



(12) 发明专利

(10) 授权公告号 CN 101986306 B

(45) 授权公告日 2013. 08. 28

(21) 申请号 201010534302. 9

CN 101329674 A, 2008. 12. 24, 全文.

(22) 申请日 2010. 11. 03

US 2010/0191768 A1, 2010. 07. 29, 全文.

(73) 专利权人 百度在线网络技术(北京)有限公司

审查员 王艳臣

地址 100085 北京市海淀区上地十街 10 号
百度大厦

(72) 发明人 兰强

(74) 专利代理机构 北京汉昊知识产权代理事务
所(普通合伙) 11370

代理人 罗朋

(51) Int. Cl.

G06F 17/30(2006. 01)

(56) 对比文件

CN 131313300 A, 2008. 11. 26, 说明书第 1 页
倒数第 5 行 - 第 31 页第 9 行, 说明书附图 1-10.

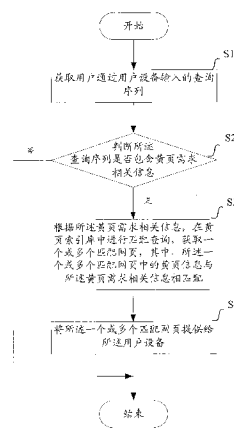
权利要求书5页 说明书23页 附图4页

(54) 发明名称

一种用于基于查询序列获取黄页信息的方法
与设备

(57) 摘要

本发明的目的是提供一种用于基于查询序列
获取黄页信息的方法与设备,其中,网络设备获取
用户通过用户设备输入的查询序列;根据所述查
询序列,判断所述查询序列是否包含黄页需求
相关信息;当所述查询序列包含黄页需求相关
信息时,根据所述黄页需求相关信息,在黄页索
引库中进行匹配查询,获取一个或多个匹配网
页,其中,所述一个或多个匹配网页中的黄页
信息与所述黄页需求相关信息相匹配;将所述
一个或多个匹配网页提供给所述用户设备。与
现有技术相比,本发明基于用户输入的查询序
列获取与之相匹配的黄页信息,不仅提高了用
户获取黄页信息的效率,也改善了用户的搜索
浏览体验。



1. 一种用于基于查询序列获取黄页信息的方法,其中,该方法包括以下步骤:
 - 步骤 a 获取用户通过用户设备输入的查询序列;
 - 步骤 b 对所述查询序列进行词义分析,当确定所述查询序列符合预定条件时,判断所述查询序列包含黄页需求相关信息,以获取所述查询序列中的全部黄页需求相关信息,其中,所述预定条件包括以下至少一项:
 - 所述查询序列包含与机构名称有关的特定类别关键词;
 - 所述查询序列包含与机构联系方式有关的特定类别关键词;
 - 所述查询序列包含与机构法人有关的特定类别关键词;
 - 步骤 c 当所述查询序列包含黄页需求相关信息时,根据所获得的所述全部黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配;
 - 步骤 d 将所述一个或多个匹配网页提供给所述用户设备。
2. 根据权利要求 1 所述的方法,其中,该方法还包括:
 - 步骤 e 将所述黄页信息以突显的形式提供给所述用户设备。
3. 根据权利要求 2 所述的方法,其中,该方法还包括:
 - 步骤 f 对所述黄页信息进行规范化处理,以获得规范化的黄页信息;其中,所述步骤 e 还包括:
 - 将所述规范化的黄页信息以突显的形式提供给所述用户设备。
4. 根据权利要求 2 或 3 所述的方法,其中,该方法还包括步骤 g:
 - 根据所述黄页需求相关信息,并结合所述查询序列,获取所述查询序列中的黄页附加需求相关信息;
 - 根据所述黄页信息与所述黄页附加需求相关信息,获取与所述黄页附加需求相关信息相匹配的黄页附加信息;其中,所述步骤 e 还包括:
 - 将所述黄页信息与所述黄页附加信息以突显的形式提供给所述用户设备。
5. 根据权利要求 2 或 3 所述的方法,其中,该方法还包括步骤 h:
 - 根据所述黄页需求相关信息,并结合所述查询序列,获取所述查询序列中的黄页焦点需求相关信息;
 - 根据所述黄页信息与所述黄页焦点需求相关信息,获取与所述黄页焦点需求相关信息相匹配的黄页焦点信息;其中,所述步骤 e 还包括:
 - 将所述黄页焦点信息以突显的形式提供给所述用户设备。
6. 根据权利要求 2 或 3 所述的方法,其中,所述突显的形式包括以下至少任一项:
 - 摘要;
 - 悬浮框;
 - 缩略图。
7. 根据权利要求 1 至 3 中任一项所述的方法,其中,所述步骤 b 还包括:
 - 根据所述查询序列,并结合该用户的历史查询操作,判断所述查询序列是否包含黄页需求相关信息。

8. 根据权利要求 1 至 3 中任一项所述的方法,其中,该方法还包括:
步骤 i 根据大量网页及其中的黄页信息,建立或更新所述黄页索引库。
9. 根据权利要求 1 至 3 中任一项所述的方法,其中,所述步骤 c 还包括:
- 当所述查询序列包含黄页需求相关信息时,根据所述黄页需求相关信息,并结合该用户的用户相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。
10. 根据权利要求 9 所述的方法,其中,所述用户相关信息包括以下至少一项:
- 用户属性;
- 用户偏好;
- 用户搜索历史记录;
- 用户设备的地址。
11. 根据权利要求 9 所述的方法,其中,该方法还包括步骤 j:
- 获取该用户对所述一个或多个匹配网页的浏览操作;
- 根据所述浏览操作,更新该用户的所述用户相关信息。
12. 根据权利要求 1 至 3 中任一项所述的方法,其中,该方法还包括:
步骤 k 根据所述黄页需求相关信息,对所述黄页需求相关信息进行预处理操作;
其中,所述步骤 c 还包括:
- 根据预处理操作后的所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。
13. 根据权利要求 12 所述的方法,其中,所述预处理操作包括以下至少一项:
- 近义词替换;
- 同义词替换;
- 词义拓展。
14. 根据权利要求 1 至 3 中任一项所述的方法,其中,该方法还包括:
步骤 l 根据预设规则,确定所述一个或多个匹配网页的优先级;
其中,所述步骤 d 还包括:
- 根据所述优先级,将所述一个或多个匹配网页提供给所述用户设备。
15. 根据权利要求 14 所述的方法,其中,所述预设规则基于包括以下至少任一项:
- 所述匹配网页的权威性;
- 所述匹配网页的累计点击量;
- 所述匹配网页的搜索排名;
- 所述匹配网页中黄页信息的更新时间。
16. 根据权利要求 1 至 3 中任一项所述的方法,其中,该方法还包括:
步骤 m 获取搜索引擎基于所述查询序列提供的第一搜索结果;
其中,所述步骤 d 还包括:
- 将所述一个或多个匹配网页和所述第一搜索结果提供给所述用户设备,其中,优先提供所述一个或多个匹配网页。
17. 根据权利要求 16 所述的方法,其中,该方法还包括:

步骤 n 将所述一个或多个匹配网页与所述第一搜索结果进行合并去冗余处理,以获得第二搜索结果;

其中,所述步骤 d 还包括:

- 将所述第二搜索结果提供给所述用户设备。

18. 一种用于基于查询序列获取黄页信息的设备,其中,该设备包括:

第一获取装置,用于获取用户通过用户设备输入的查询序列;

判断装置,用于对所述查询序列进行词义分析,当确定所述查询序列符合预定条件时,判断所述查询序列包含黄页需求相关信息,以获取所述查询序列中的全部黄页需求相关信息,其中,所述预定条件包括以下至少一项:

- 所述查询序列包含与机构名称有关的特定类别关键词;

- 所述查询序列包含与机构联系方式有关的特定类别关键词;

- 所述查询序列包含与机构法人有关的特定类别关键词;

查询装置,用于当所述查询序列包含黄页需求相关信息时,根据所获得的所述全部黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配;

第一提供装置,用于将所述一个或多个匹配网页提供给所述用户设备。

19. 根据权利要求 18 所述的设备,其中,该设备还包括:

第二提供装置,用于将所述黄页信息以突显的形式提供给所述用户设备。

20. 根据权利要求 19 所述的设备,其中,该设备还包括:

规范化处理装置,用于对所述黄页信息进行规范化处理,以获得规范化的黄页信息;

其中,所述第二提供装置还用于:

- 将所述规范化的黄页信息以突显的形式提供给所述用户设备。

21. 根据权利要求 19 或 20 所述的设备,其中,该设备还包括第二获取装置,用于:

- 根据所述黄页需求相关信息,并结合所述查询序列,获取所述查询序列中的黄页附加需求相关信息;

- 根据所述黄页信息与所述黄页附加需求相关信息,获取与所述黄页附加需求相关信息相匹配的黄页附加信息;

其中,所述第二提供装置还用于:

- 将所述黄页信息与所述黄页附加信息以突显的形式提供给所述用户设备。

22. 根据权利要求 19 或 20 所述的设备,其中,该设备还包括第三获取装置,用于:

- 根据所述黄页需求相关信息,并结合所述查询序列,获取所述查询序列中的黄页焦点需求相关信息;

- 根据所述黄页信息与所述黄页焦点需求相关信息,获取与所述黄页焦点需求相关信息相匹配的黄页焦点信息;

其中,所述第二提供装置还用于:

- 将所述黄页焦点信息以突显的形式提供给所述用户设备。

23. 根据权利要求 19 或 20 所述的设备,其中,所述突显的形式包括以下至少任一项:

- 摘要;

- 悬浮框;

- 缩略图。

24. 根据权利要求 18 至 20 中任一项所述的设备,其中,所述判断装置还用于:

- 根据所述查询序列,并结合该用户的历史查询操作,判断所述查询序列是否包含黄页需求相关信息。

25. 根据权利要求 18 至 20 中任一项所述的设备,其中,该设备还包括:

第一更新装置,用于根据大量网页及其中的黄页信息,建立或更新所述黄页索引库。

26. 根据权利要求 18 至 20 中任一项所述的设备,其中,所述查询装置还用于:

- 当所述查询序列包含黄页需求相关信息时,根据所述黄页需求相关信息,并结合该用户的用户相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。

27. 根据权利要求 26 所述的设备,其中,所述用户相关信息包括以下至少一项:

- 用户属性;
- 用户偏好;
- 用户搜索历史记录;
- 用户设备的地址。

28. 根据权利要求 25 所述的设备,其中,该设备还包括第二更新装置,用于:

- 获取该用户对所述一个或多个匹配网页的浏览操作;
- 根据所述浏览操作,更新该用户的所述用户相关信息。

29. 根据权利要求 18 至 20 中任一项所述的设备,其中,该设备还包括:

预处理装置,用于根据所述黄页需求相关信息,对所述黄页需求相关信息进行预处理操作;

其中,所述查询装置还用于:

- 根据预处理操作后的所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。

30. 根据权利要求 29 所述的设备,其中,所述预处理操作包括以下至少一项:

- 近义词替换;
- 同义词替换;
- 词义拓展。

31. 根据权利要求 18 至 20 中任一项所述的设备,其中,该设备还包括:

排序装置,用于根据预设规则,确定所述一个或多个匹配网页的优先级;

其中,所述第一提供装置还用于:

- 根据所述优先级,将所述一个或多个匹配网页提供给所述用户设备。

32. 根据权利要求 31 所述的设备,其中,所述预设规则基于包括以下至少任一项:

- 所述匹配网页的权威性;
- 所述匹配网页的累计点击量;
- 所述匹配网页的搜索排名;
- 所述匹配网页中黄页信息的更新时间。

33. 根据权利要求 18 至 20 中任一项所述的设备,其中,该设备还包括:

第四获取装置,用于获取搜索引擎基于所述查询序列提供的第一搜索结果;

其中,所述第一提供装置还用于:

- 将所述一个或多个匹配网页和所述第一搜索结果提供给所述用户设备,其中,优先提供所述一个或多个匹配网页。

34. 根据权利要求 33 所述的设备,其中,该设备还包括:

合并装置,用于将所述一个或多个匹配网页与所述第一搜索结果进行合并去冗余处理,以获得第二搜索结果;

其中,所述第一提供装置还用于:

- 将所述第二搜索结果提供给所述用户设备。

35. 一种搜索引擎,包括如权利要求 18 至 34 中任一项所述的基于查询序列获取黄页信息的设备。

一种用于基于查询序列获取黄页信息的方法与设备

技术领域

[0001] 本发明涉及互联网技术领域,尤其涉及一种用于基于查询序列获取黄页信息的技术。

背景技术

[0002] 随着互联网应用的普及推广和对人们生活的渗透,越来越多的用户开始习惯于从互联网获取其所需的信息,例如机构的地址、电话、传真、联系人、邮箱等黄页信息。

[0003] 然而,随着人机交互等计算机技术的进一步发展,人们从主动获取资源开始转为依赖于互联网主动提供所需的资源。例如,用户在查找某些机构的黄页信息时,其并不知道哪些网页中包含有所想要查找的黄页信息。

[0004] 因此,如何基于用户输入的查询序列在全互联网范围内向其提供黄页信息,以提高用户高效获取信息的搜索浏览体验,成为现今亟需解决的一个问题。

发明内容

[0005] 本发明的目的是提供一种用于基于查询序列获取黄页信息的方法与设备。

[0006] 根据本发明的一个方面,提供了一种用于基于查询序列获取黄页信息的方法,该方法包括以下步骤:

[0007] a 获取用户通过用户设备输入的查询序列;

[0008] b 根据所述查询序列,判断所述查询序列是否包含黄页需求相关信息;

[0009] c 当所述查询序列包含黄页需求相关信息时,根据所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配;

[0010] d 将所述一个或多个匹配网页提供给所述用户设备。

[0011] 根据本发明的另一方面,还提供了一种用于基于查询序列获取黄页信息的设备,该设备包括:

[0012] 第一获取装置,用于获取用户通过用户设备输入的查询序列;

[0013] 判断装置,用于根据所述查询序列,判断所述查询序列是否包含黄页需求相关信息;

[0014] 查询装置,用于当所述查询序列包含黄页需求相关信息时,根据所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配;

[0015] 第一提供装置,用于将所述一个或多个匹配网页提供给所述用户设备。

[0016] 根据本发明的又一方面,还提供了一种搜索引擎,包括上述基于查询序列获取黄页信息的装置。

[0017] 与现有技术相比,本发明基于用户输入的查询序列获取与之相匹配的黄页信息,不仅提高了用户获取黄页信息的效率,也改善了用户的搜索浏览体验。

附图说明

[0018] 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述,本发明的其它特征、目的和优点将会变得更明显:

[0019] 图 1 示出根据本发明一个方面的用于基于查询序列获取黄页信息的系统拓扑图;

[0020] 图 2 示出根据本发明一个方面的用于基于查询序列获取黄页信息的设备示意图;

[0021] 图 3 示出根据本发明一个优选实施例的用于基于查询序列获取黄页信息的设备示意图;

[0022] 图 4 示出根据本发明另一个方面的用于基于查询序列获取黄页信息的方法流程图;

[0023] 图 5 示出根据本发明一个优选实施例的用于基于查询序列获取黄页信息的方法流程图。

[0024] 附图中相同或相似的附图标记代表相同或相似的部件。

具体实施方式

[0025] 下面结合附图对本发明作进一步详细描述。

[0026] 图 1 示出根据本发明一个方面的用于基于查询序列获取黄页信息的系统拓扑图,其中包括网络设备 1 以及多个与之经由网络相连接的用户设备 2。其中,网络设备 1 可与多个用户设备 2 经由网络相连接,并获取用户通过用户设备 2 输入的查询序列;根据所述查询序列,判断所述查询序列是否包含黄页需求相关信息;当所述查询序列包含黄页需求相关信息时,根据所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配;将所述一个或多个匹配网页提供给所述用户设备。在此,网络包括但不限于互联网、广域网、城域网、局域网、VPN 网络、无线自组织网络(Ad Hoc 网络)等。

[0027] 另外,网络设备 1 包括但不限于网络主机、单个网络服务器、多个网络服务器集或多个服务器构成的云。在此,云是基于云计算(Cloud Computing)的大量计算机或网络服务器构成,其中,云计算是分布式计算的一种,由一群松散耦合的计算机集组成的一个超级虚拟计算机。而用户设备 2 包括但不限于任何一种可与用户通过键盘、鼠标、遥控器、触摸板、或手写设备等方式进行人机交互的电子产品,例如计算机、手机、PDA、掌上电脑 PPC 或 IPTV 等。网络设备 1 与用户设备 2 之间的通信相互独立,可以是基于诸如 TCP/IP 协议、UDP 协议等的分组数据传输。

[0028] 本领域技术人员应能理解上述网络设备 1、用户设备 2 以及连接其间的网络或通信方式仅为举例,其他现有的或今后可能出现的网络设备、用户设备或网络、通信方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0029] 本领域技术人员还应能理解,图 1 中仅为简明起见而示出的各类网络元素的数量可能小于一个实际网络中的数量,但这种省略无疑地是以不会影响对本发明进行清楚、充分的公开为前提的。

[0030] 为简明起见,下面以一个用户设备 2 为例进行描述。本领域技术人员应能理解,网络设备 1 可以同时与多个用户设备 2 交互,并根据从不同用户设备 2 输入的查询序列;根据

所述查询序列,判断所述查询序列是否包含黄页需求相关信息;当所述查询序列包含黄页需求相关信息时,根据所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配;将所述一个或多个匹配网页提供给与之相对应的用户设备 2。

[0031] 图 2 示出根据本发明一个方面的基于查询序列获取黄页信息的设备示意图,其中,网络设备 1 包括第一获取装置 11、判断装置 12、查询装置 13 和第一提供装置 14。

[0032] 其中,第一获取装置 11 用于获取用户通过用户设备输入的查询序列。具体地,用户通过与用户设备 2 的交互方式,包括但不限于键盘、鼠标、遥控器、触摸板、或手写设备,在浏览器软件、应用程序或客户端软件中输入查询序列;第一获取装置 11 通过页面技术,例如 JSP、ASP、PHP 技术,实时地获取用户输入的查询序列。本领域技术人员应能理解上述页面技术仅为举例,其他现有的或今后可能出现的页面技术如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。本领域技术人员还应能理解上述获取用户浏览行为的技术仅为举例,其他现有的或今后可能出现的获取用户浏览行为的技术如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0033] 判断装置 12 用于根据所述查询序列,判断所述查询序列是否包含黄页需求相关信息。具体地,判断装置 12 根据第一获取装置 11 获取的用户输入的查询序列,通过例如词义分析、查询用户历史查询操作等方式,判断所述查询序列是否包含黄页需求相关信息。以词义分析为例,所述查询序列中可能包含多个黄页需求相关信息,当判断装置 12 检测到其中一个黄页需求相关信息时,即可以判断所述查询序列包含黄页需求相关信息,但此后,判断装置 12 仍需继续执行操作,以获取所述查询序列中的全部黄页需求相关信息,以方便查询装置 13 根据所述全部黄页需求相关信息进行匹配查询。优选地,所述判断装置 12 也可以先进行获取的操作,获取所述查询序列中的全部黄页需求相关信息,再判断有无获取到所述黄页需求相关信息,若有,则查询装置 13 继续执行下一步操作。例如,第一获取装置 11 获取到用户输入的查询序列为“百度地址”,判断装置 12 检测到该查询序列中包含的“百度”为黄页需求相关信息,判断该查询序列包含黄页需求相关信息,此后,判断装置 12 继续进行检测,获取该查询序列中的另一个黄页需求相关信息“地址”。又如,判断装置 12 可以获取该查询序列中的全部黄页需求相关信息“百度”和“地址”,再判断该查询序列中包含有黄页需求相关信息。本领域技术人员应能理解上述判断查询序列是否包含黄页需求相关信息的方式仅为举例,其他现有的或今后可能出现的判断查询序列是否包含黄页需求相关信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0034] 查询装置 13 用于当所述查询序列包含黄页需求相关信息时,根据所述黄页需求相关信息,在黄页索引库 15 中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。具体地,当判断装置 12 判断得出所述查询序列包含黄页需求相关信息时,查询装置 13 根据所述黄页需求相关信息,在黄页索引库 15 中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。在此,所述黄页索引库 15 中存储有大量包含有黄页信息网页,例如,网络设备 1 可以根据大量的网页和其中的黄页信息,通过建立或更新数据表等形式,建立或更新所述黄页索引库 15,在所述数据表中,可以是一个黄页信

息对应一个网页,也可以是一个黄页信息对应包含有该黄页信息的几个网页。查询装置 13 根据所述黄页需求相关信息,在黄页索引库 15 中进行匹配查询,首先获取与所述黄页需求相关信息相匹配的黄页信息,再根据所述黄页信息,匹配获取一个或多个与所述黄页信息相对应的匹配网页。例如,在黄页索引库 15 中,存储有“百度时代网络技术(北京)有限公司”、“百度在线网络技术有限公司”、“百度时代网络技术(北京)有限公司客户发展部”三个标题的网页及其中的黄页信息,用户输入查询序列“百度北京”,查询装置 13 根据其中的黄页需求相关信息“百度北京”,在黄页索引库 15 中匹配到与之相对应的上述三个网页的黄页信息,并根据所述黄页信息,匹配获取到上述三个网页。本领域技术人员应能理解上述获取匹配网页的方式仅为举例,其他现有的或今后可能出现的获取匹配网页的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0035] 第一提供装置 14 用于将所述一个或多个匹配网页提供给所述用户设备 2。具体地,第一提供装置 14 根据查询装置 13 获取的一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配,通过页面技术,例如 JSP、ASP、PHP 技术,将所述一个或多个匹配网页以链接、摘要或缩略图等形式,按一定顺序和格式,嵌入在搜索结果页面中,提供给所述用户设备 2,供用户进行浏览。接前例,第一提供装置 14 将查询装置 13 获取的“百度时代网络技术(北京)有限公司”、“百度在线网络技术有限公司”、“百度时代网络技术(北京)有限公司客户发展部”三个标题的网页,以链接、摘要或缩略图等形式,按一定顺序和格式,嵌入于搜索结果页面之中或之上,提供给所述用户设备 2。本领域技术人员应能理解上述将匹配网页提供给用户设备的方式仅为举例,其他现有的或今后可能出现的将推荐信息提供给用户设备的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0036] 优选地,网络设备 1 的各个装置之间是持续不断工作的。具体地,第一获取装置 11 获取用户通过用户设备 2 输入的查询序列;随后,判断装置 12 根据所述查询序列,判断所述查询序列是否包含黄页需求相关信息;当所述查询序列包含黄页需求相关信息时,查询装置 13 根据所述黄页需求相关信息,在黄页索引库 15 中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配;接着,第一提供装置 14 再将所述一个或多个匹配网页提供给所述用户设备。在此,本领域技术人员应理解“持续”是指网络设备 1 的各装置分别按照设定的或实时调整的工作模式要求进行查询序列的获取、黄页需求相关信息的判断、匹配网页的获取及提供,直至该用户在较长时间内停止通过用户设备 2 输入查询序列。

[0037] 优选地,网络设备 1 还可以包括第二提供装置(未示出)。所述第二提供装置用于将所述黄页信息以突显的形式提供给所述用户设备。具体地,第二提供装置根据查询装置 13 获取的一个或多个匹配网页中的黄页信息,通过页面技术,例如 JSP、ASP、PHP 技术,将所述黄页信息以摘要、悬浮框或缩略图等形式,按一定顺序和格式,嵌入于搜索结果页面之中或之上,提供给所述用户设备 2,供用户进行浏览。在此,黄页信息的展现应与前述匹配网页的展现相对应,例如当所述黄页信息以摘要的形式展现在搜索结果页面时,可以在所述匹配网页的链接的下方展现给用户。优选地,还可以将所述黄页信息加粗、加黑,或改变字体、字形、颜色,或增加背景颜色等方式,将其进行突显展现。当以悬浮框的形式提供给用户设备时,可以是在页面的某个固定的区域显示,也可以由用户设定或根据浏览对象的位置确

定,其窗口大小可以进行调节。例如,将所述黄页信息以悬浮框的形式显示在用户设备 2 的输入框中,其可以是透明的以不影响用户的输入,并随着用户的输入操作不断变换其所处的位置,例如随着输入序列变长,所述悬浮框不断右移。优选地,所述悬浮框还可以根据光标与之的距离自动调整自身的透明度,以不影响用户的正常浏览;当光标移动到距所述悬浮框的设定距离范围外时隐藏所述悬浮框;当在所述悬浮框之外点击时,隐藏所述悬浮框。本领域技术人员应能理解上述提供黄页信息的方式仅为举例,其他现有的或今后可能出现的提供黄页信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0038] 本领域技术人员应能理解,上述第一提供装置、第二提供装置仅为示例,在实践中,它们可以是两个独立的模块,也可全部集成在一个模块中。

[0039] 优选地,网络设备 1 还可以包括规范化处理装置(未示出)。所述规范化处理装置用于对所述黄页信息进行规范化处理,以获得规范化的黄页信息;其中,所述第二提供装置还用于将所述规范化的黄页信息以突显的形式提供给所述用户设备 2。查询装置 13 匹配查询到的匹配网页中的黄页信息可能存在多种问题:1) 形式上并不统一,例如有的网页在电话号码前加上了区号,有的没有加区号,再如有的匹配网页中的黄页信息包含电子邮箱地址,有的则不包含;2) 表述上不规范,如将“中国人民大学附属中学”记为“人大附中”。规范化处理装置可以将所述黄页信息进行规范化处理,以获得形式统一、表述规范的黄页信息。例如,规范化处理装置根据预置的黄页模板,对所述黄页信息进行规范化处理,例如将电话号码统一按照“国际区号-国内区号-电话号码”的形式显示,又如假设规定提供给用户的黄页信息包含有项目“机构名称、地址、邮编、网址、电话、传真、法定代表人”并按照一定的格式排版,若该网页中不包含某一项的信息,则相应的项目空置。又如,规范化处理装置可以根据对应于同一个黄页需求相关信息的多个匹配网页中的黄页信息,通过归类去重处理得到对应于该黄页需求相关信息的黄页模板,然后根据此黄页模板对这多个匹配黄页中的行黄页信息进行规范化,如将这些黄页信息按归类去重处理所得到的黄页模板进行格式排版,将该黄页信息中的信息项填入该黄页模板的对应项目中,若该黄页信息不包含某个信息项,则空置该黄页模板中的对应项目。再如,该规范化处理装置还可以通过自然语言处理技术,使得黄页信息中的信息项表述更为规范,例如该规范处理装置利用自然语言处理技术,将机构名称按 Term 切分,诸如“人大附中”,可分割为“人大”、“附中”;并结合同义词列表、别名列表,可将其规范化为“中国人民大学附属中学”。此后,第二提供装置再通过页面技术,例如 JSP、ASP、PHP 技术,将所述模板嵌入于返回给用户设备的搜索结果页面之中或之上,以突显的形式提供给用户设备 2。本领域技术人员应能理解,该规范化处理装置还可以通过黄页模板与自然语言处理技术的结合,获取形式统一且表述规范的黄页信息。本领域技术人员还应能理解上述规范化处理黄页信息的方式仅为举例,其他现有的或今后可能出现的规范化处理黄页信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0040] 优选地,网络设备 1 还可以包括第二获取装置(未示出)。所述第二获取装置用于根据所述黄页需求相关信息,并结合所述查询序列,获取所述查询序列中的黄页附加需求相关信息;根据所述黄页信息与所述黄页附加需求相关信息,获取与所述黄页附加需求相关信息相匹配的黄页附加信息;其中,所述第二提供装置还用于将所述黄页信息与所述黄

页附加信息以突显的形式提供给所述用户设备 2。在用户输入的查询序列中,除与机构名称有关、与机构联系方式有关或与机构法人有关的黄页需求相关信息外,还可能包含黄页附加需求相关信息,例如与该机构有关的地图相关信息、路线相关信息等需求。具体地,第二获取装置根据判断装置 12 所获取的黄页需求相关信息,并结合第一获取装置 11 获取的查询序列,通过词义分析、数据库查询等方式,获取所述查询序列中的黄页附加需求相关信息;此后,第二获取装置根据所述黄页信息与所述黄页附加需求相关信息,通过数据库查询或利用第三方设备提供的应用程序接口(API)从第三方设备获取等方式,获取与所述黄页附加需求相关信息相匹配的黄页附加信息;随后,第二提供装置通过页面技术,例如 JSP、ASP、PHP 技术,将所述黄页信息与所述黄页附加信息嵌入于返回给用户设备 2 的搜索结果页面之中或之上,以突显的形式提供给用户设备 2。例如,用户通过用户设备 2 输入查询序列“百度北京怎么去”,判断装置 12 通过词义分析,获取到该用户的黄页需求相关信息为“百度北京”,第二获取装置根据该用户的黄页需求相关信息,并结合该用户输入的查询序列,通过词义分析,获取到黄页附加需求相关信息为“地图”(或“乘车路线”);此后,第二获取装置再根据“百度北京”的黄页信息与所述黄页附加需求相关信息,通过调用第三方地图搜索设备获取“百度北京”的地图信息(或者通过数据库查询获得“百度北京”周边的乘车信息,如轨道交通与公交线路);随后,第二提供装置再将所述“百度北京”的黄页信息与地图信息(或乘车路线)提供给用户设备 2。优选地,第二获取装置还可以取用户输入的查询序列所包括的另一位置信息或者用户当前的定位信息(假设用户设备 2 包含有 GPS 等地理信息模块等)作为一个端点,提供该端点与所述黄页信息所对应的另一个端点之间的详细的路线图。本领域技术人员应能理解上述获取与提供黄页附加信息的方式仅为举例,其他现有的或今后可能出现的获取或提供黄页附加信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0041] 优选地,网络设备 1 还可以包括第三获取装置(未示出)。所述第三获取装置用于根据所述黄页需求相关信息,并结合所述查询序列,获取所述查询序列中的黄页焦点需求相关信息;根据所述黄页信息与所述黄页焦点需求相关信息,获取与所述黄页焦点需求相关信息相匹配的黄页焦点信息;此后,第二提供装置将所述黄页焦点信息以突显的形式提供给所述用户设备。由于对于某一部分用户来讲,其可能只需要黄页信息中的黄页焦点信息,例如地址、电话、网站中的一个或多个,网络设备 1 需要获取该用户的黄页焦点需求相关信息,并进一步获取与之相匹配的黄页焦点信息。具体地,第三获取装置根据判断装置 12 所获取的黄页需求相关信息,并结合第一获取装置 11 获取的查询序列,获取所述查询序列中的黄页焦点需求相关信息;此后,第三获取装置根据所述黄页信息与所述黄页焦点需求相关信息,通过词义分析、数据库查询等方式,获取与所述黄页焦点需求相关信息相匹配的黄页焦点信息;随后,第二提供装置通过页面技术,例如 JSP、ASP、PHP 技术,将所述黄页焦点信息嵌入于返回给用户设备 2 的搜索结果页面之中或之上,以突显的形式提供给用户设备 2。例如,用户通过用户设备 2 输入查询序列“联想客服电话”,判断装置 12 通过词义分析,获取到该用户的黄页需求相关信息为“联想 lenovo”,第三获取装置根据该用户的黄页需求相关信息,并结合该用户输入的查询序列,通过词义分析,获取到黄页焦点需求相关信息为“客服电话”;此后,第三获取装置再通过词义分析、数据库查询等方式,获取与所述黄页焦点需求相关信息相匹配的黄页焦点信息“8008108888”;随后,第二提供装置通过页

面技术,例如 JSP、ASP、PHP 技术,将所述黄页焦点信息“8008108888”嵌入于返回给用户设备 2 的搜索结果页面之中或之上,以突显的形式提供给用户设备 2。例如,将所述黄页焦点信息“8008108888”以悬浮框的形式显示在用户设备 2 的输入框中,其可以是透明的以不影响用户的输入,并随着用户的输入操作不断变换其所处的位置,例如随着输入序列变长,所述悬浮框不断右移。本领域技术人员应能理解上述获取与提供黄页焦点信息的方式仅为举例,其他现有的或今后可能出现的获取或提供黄页焦点信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0042] 优选地,第二提供装置将黄页信息提供给用户设备 2 的突显方式,包括但不限于,以下至少一项:摘要、悬浮框和缩略图等形式。第二提供装置可以通过页面技术,例如 JSP、ASP、PHP 技术,将查询装置 13 获取的黄页信息以摘要的形式提供给所述用户设备 2,黄页信息的展现应与该黄页信息所属的匹配网页的展现相对应,可以在所述匹配网页的链接的下方展现给用户。优选地,还可以将所述黄页信息加粗、加黑,或改变字体、字形、颜色,或增加背景颜色等方式,将其进行突显展示。第二提供装置还可以将查询装置 13 获取的黄页信息,以悬浮框的形式提供给所述用户设备 2,用户通过点击网页上特定的按钮或用户自定义的按钮或按键,或者将光标悬停指向搜索结果页面中的链接时,打开所述悬浮框,以浏览所述黄页信息。优选地,所述悬浮框可以是在页面的某个固定的区域显示,也可以由用户设定或根据浏览对象的位置确定,其窗口大小可以进行调节。所述悬浮框还可以根据光标与之的距离自动调整自身的透明度,以不影响用户的正常浏览;当光标移动到距所述悬浮框的设定距离范围外时隐藏所述悬浮框;当在所述悬浮框之外点击时,隐藏所述悬浮框。例如,将所述黄页信息以悬浮框的形式显示在用户设备 2 的输入框中,其可以是透明的以不影响用户的输入,并随着用户的输入操作不断变换其所处的位置,例如随着输入序列变长,所述悬浮框不断右移。第二提供装置还可以将查询装置 13 获取的黄页信息,以缩略图的形式提供给所述用户设备 2,用户可以通过点击特定的或自定义的按钮放大或缩小所述缩略图,以浏览所述黄页信息。在此,本领域技术人员应能理解,上述突显的形式同样适用于第二提供装置将规范化的黄页信息、黄页信息与黄页附加信息以突显的形式提供给所述用户设备 2 本领域技术人员还应能理解上述突显的形式仅为举例,其他现有的或今后可能出现的突显的形式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0043] 优选地,判断装置 12 还用于对所述查询序列进行词义分析,当确定所述查询序列符合预定条件时,判断所述查询序列包含黄页需求相关信息,其中,所述预定条件包括但不限于以下至少一项:a) 所述查询序列包含与机构名称有关的特定类别关键词;b) 所述查询序列包含与机构联系方式有关的特定类别关键词;c) 所述查询序列包含与机构法人有关的特定类别关键词。具体地,判断装置 12 根据第一获取装置 11 获取的用户输入的查询序列,通过词义分析,判断所述查询序列中是否包含与机构名称有关的特定类别关键词,例如机构的部分名称或全部名称,是否包含与机构联系方式有关的特定类别关键词,例如电话、邮箱、传真、网址等,是否包含与机构法人有关的特定类别关键词,例如法人代表、法定代表人等;当所述查询序列包含上述特定类别关键词时,判断所述查询序列包含黄页需求相关信息。本领域技术人员应能理解上述判断查询序列是否包含黄页需求相关信息的预定条件仅为举例,其他现有的或今后可能出现的判断查询序列是否包含黄页需求相关信息的预定条件如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0044] 更优选地,判断装置 12 还用于根据所述查询序列,并结合该用户的历史查询操作,判断所述查询序列是否包含黄页需求相关信息。具体的,判断装置 12 根据第一获取装置 11 获取的用户输入的查询序列,并结合该用户的历史查询操作,如用户曾使用的查询关键词、用户对包含黄页信息的网页的点击浏览记录,判断所述查询序列是否包含黄页需求相关信息。例如,判断装置 12 根据用户当前的查询序列为“百度”,结合该用户的历史查询操作,如曾经点击包含有“百度公司”黄页信息的网页的 URL,则判断该查询序列包括黄页需求相关信息。又如,判断装置 12 根据用户当前的查询序列为“百度”,结合该用户的历史查询操作,如该用户曾经变换查询序列,例如用户输入查询序列“百度”,在获得的搜索结果页面中点击联想查询序列“百度公司的地址”,后者为明确的黄页需求,则判断该查询序列包括黄页需求相关信息。判断装置 12 还可以对所述查询序列进行词义分析,并结合该用户的历史查询操作,当确定该查询序列满足预定条件时,判断该查询序列包含黄页需求相关信息。例如,判断装置 12 根据用户当前的查询序列为“百度”,结合该用户的历史查询操作,如该用户曾经变换查询序列,例如用户输入查询序列“百度”,在获得的搜索结果页面中点击联想查询序列“百度公司法人代表”,后者为明确的黄页需求,则判断该查询序列包括黄页需求相关信息。本领域技术人员还应能理解上述判断查询序列是否包含黄页需求相关信息的方式仅为举例,其他现有的或今后可能出现的判断查询序列是否包含黄页需求相关信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。在此,该用户的所述历史查询操作既可以是网络设备 1 独立获取的,也可以是接受来自其他产品或第三方设备所完成的对该用户的所述历史查询操作的获取。例如,可通过检测该用户使用网络设备 1 的搜索服务过程中的操作,或通过向用户设备 2 询问并获取该用户的 COOKIE 信息来获得该用户的历史查询操作,优选地,可赋予该用户一个临时 ID 并为其建立一个历史查询操作库,将此次搜索过程中用户的历史记录保存于该临时历史查询操作库中。本领域技术人员还应能理解上述获取用户的历史查询操作的方式仅为举例,其他现有的或今后可能出现的获取用户的历史查询操作的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0045] 优选地,网络设备 1 还可以包括第一更新装置(未示出)。所述第一更新装置用于根据大量网页及其中的黄页信息,建立或更新所述黄页索引库 15。具体地,第一更新装置对通过爬虫工具等网页抓取工具直接抓取的大量网页,或者从第三方设备中获取的存储于该第三方设备的大量网页,进行特征分析,以获取包含黄页信息的潜在目标页,并根据这些潜在目标页及其中的黄页信息,建立或更新所述黄页索引库 15。在此,第一更新装置确定潜在目标页的方式,包括但不限于:1) 根据网址中包含有例如“show_company”、“huangye”、“yellowpage”等表意信息的网页为潜在目标页;2) 根据网页 Title 和 Meta 信息域中含有“公司、企业、公司介绍、企业简介”等表意信息的网页为潜在目标页;3) 根据在网页原文中提取有“地址、电话、传真、邮编、联系人、名称”等表意的信息对的网页为潜在目标页;4) 或者根据其它的指向信息,例如 anchor(锚文本)信息中包含有以上表意信息等的网页为潜在目标页;5) 上述几种方式的任意组合。第一更新装置通过相互补充验证,提取潜在目标页中的有效机构名称、地址信息、邮编、电话、联系人、传真等黄页信息。将所述黄页信息与所述潜在目标页通过建立或更新数据表等形式,建立或更新所述黄页索引库 15。在所述数据表中,可以是一个黄页信息对应一个网页,也可以是一个黄页信息对应包含有该黄页信

息的几个网页。例如,在黄页索引库 15 中,“百度在线网络技术有限公司”的黄页信息“公司全称:百度在线网络技术有限公司;网址:http://www.baidu.com;电话:86-10-82621188;电子邮件:webmaster@baidu.com”可以只对应“百度在线网络技术有限公司”一个网页,也可以同时对应“百度在线网络技术有限公司_百度在线网络技术有限公司简介_艾瑞网行业公司”这个网页。本领域技术人员应能理解上述建立或更新黄页索引库 15 的方式仅为举例,其他现有的或今后可能出现的建立或更新黄页索引库 15 的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0046] 优选地,查询装置 13 还用于当所述查询序列包含黄页需求相关信息时,根据所述黄页需求相关信息,并结合该用户的用户相关信息,在黄页索引库 15 中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。具体地,当判断装置 12 判断所述查询序列包含黄页需求相关信息时,查询装置 13 根据所述黄页需求相关信息,并结合该用户的用户相关信息,如该用户的搜索历史记录或用户设备的地址,在黄页索引库 15 中进行匹配查询,以获得与用户实际需求相匹配的一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。例如,查询装置 13 先将用户相关信息的关键词与黄页需求相关信息组合成第二查询序列,随后,查询装置 13 再根据所述第二查询序列,在黄页索引库 15 中进行查询匹配,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。在此,所述用户相关信息既可以是网络设备 1 独立获取的,也可以是接受来自用户设备 2 记录的,或者接受来自其他产品或第三方设备所获取的用户相关信息。本领域技术人员应能理解上述获取用户相关信息的方法仅为举例,其他现有的或今后可能出现的获取用户相关信息的方法如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0047] 优选地,所述用户相关信息包括但不限于以下至少一项:a) 用户属性;b) 用户偏好;c) 用户搜索历史记录;d) 用户设备的地址。用户属性包括但不限于,用户预先输入的个人背景信息,例如性别、年龄、职业、宗教、国籍、教育程度等。当所述查询序列包含黄页需求相关信息时,查询装置 13 根据所述黄页需求相关信息,并结合该用户的用户属性,在黄页索引库 15 中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。例如,用户输入查询序列“新东方”进行查询,当用户为学生时,提供有关“新东方学校”的黄页信息及相应的匹配网页,当用户为厨师时,提供有关“新东方厨师学校”的黄页信息及相应的匹配网页。用户偏好包括但不限于,用户预先设置的一些偏好选择,例如关注的机构类型、行业信息等。当所述查询序列包含黄页需求相关信息时,查询装置 13 根据所述黄页需求相关信息,并结合该用户的用户偏好,在黄页索引库 15 中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。例如,用户设置其偏好的机构类型为客服公司,则当用户输入查询序列“百度北京”时,优先提供有关“百度公司客服”的黄页信息及相应的网页,当用户设置其偏好的机构类型为研发公司时,其输入查询序列“百度北京”,优先提供有关“百度在线网络技术有限公司(北京)有限公司”等的黄页信息及相应的网页。用户搜索历史记录,其由系统来记录和统计用户先前的搜索和点击浏览历史,从而可判断或预测用户的需求。其中,当用户为登录用户时,网络设备 1 可通过该用户的 ID 在网络端保

存有该用户相关信息的用户特征库中进行查询；当用户为非登录用户时，可通过检测该用户使用本网络设备 1 的搜索服务过程中的动作，或通过向用户设备 2 询问并获取该用户的 COOKIE 信息来获得，优选地，可赋予该用户一个临时 ID 并为其建立一个用户特征库，将此次搜索过程中用户的历史记录保存于该临时用户特征库中。例如，用户输入查询序列“复旦大学”，提供的匹配网页可能包括关于“复旦大学”、“复旦大学医学院”、“复旦大学出版社”等的黄页信息，当用户点击“复旦大学医学院”时，网络设备 1 或用户设备 2 记录下用户的此次浏览操作记录，当用户再次输入“复旦大学”或其他大学的名称作为查询序列时，网络设备 1 优先将关于“某某大学出版社”的黄页信息及相应的匹配网页提供给该用户。用户设备的地址，其可以反映该用户当前所处地域，可由网络设备 1 实时检测用户设备 2 的 IP 地址来获取。例如，用户通过用户设备 2 输入查询序列“新东方语言学校”，网络设备 1 检测到该用户设备 2 的 IP 地址位于北京，则优先将北京新东方语言学校的黄页信息及相应的匹配网页提供给该用户。本领域技术人员应能理解上述用户相关信息仅为举例，其他现有的或今后可能出现的用户相关信息如可适用于本发明，也应包含在本发明保护范围以内，并在此以引用方式包含于此。

[0048] 优选地，网络设备 1 还包括第二更新装置（未示出）。所述第二更新装置用于获取该用户对所述一个或多个匹配网页的浏览操作；根据所述浏览操作，更新该用户的所述用户相关信息。具体地，所述第二更新装置根据与用户设备 2 及用户的交互，例如通过 JSP、ASP、PHP 等页面技术，获取该用户对所述一个或多个匹配网页的浏览操作；并根据这些浏览操作更新该用户的用户相关信息。在此，所述第二更新装置更新用户相关信息的方式包括但不限于：1) 第二更新装置根据用户更改的预设信息，更新该用户的用户相关信息；2) 第二更新装置可在用户一次浏览操作结束后记录用户在此浏览过程中的所有动作，例如点击、浏览或关闭；3) 第二更新装置在用户一次浏览过程中记录用户每次动作并实时对网络端该用户的特征库进行更新，从而可在用户在此次浏览过程中点击其他黄页信息时提供最新的用户相关信息。例如，用户输入“复旦大学”，提供的匹配网页可能包括关于“复旦大学”、“复旦大学医学院”、“复旦大学出版社”等的黄页信息，当用户点击“复旦大学医学院”时，记录该用户的浏览操作并更新该用户的历史记录。本领域技术人员应能理解上述更新用户相关信息的方式仅为举例，其他现有的或今后可能出现的更新用户相关信息的方式如可适用于本发明，也应包含在本发明保护范围以内，并在此以引用方式包含于此。

[0049] 本领域技术人员应能理解，上述第一更新装置、第二更新装置仅为示例，在实践中，它们可以是两个独立的模块，也可全部集成在一个模块中。

[0050] 图 3 示出根据本发明一个优选实施例的用于基于查询序列获取黄页信息的设备示意图，网络设备 1 还可以包括预处理装置 16'。其中，第一获取装置 11'、判断装置 12'、查询装置 13' 和第一提供装置 14' 分别与图 2 所示对应装置 11-14 相同，故此处不再赘述，并通过引用的方式包含于此。

[0051] 其中，预处理装置 16' 用于根据所述黄页需求相关信息，对所述黄页需求相关信息进行预处理操作；其中，查询装置 13' 根据预处理操作后的所述黄页需求相关信息，在黄页索引库 15' 中进行匹配查询，获取一个或多个匹配网页，其中，所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。具体地，预处理装置 16' 根据判断装置 12' 所获取的黄页需求相关信息，通过自然语言处理技术，例如近义词替换、同义词替换、词义

拓展等方式,对所述黄页需求相关信息进行预处理操作;随后,查询装置 13' 根据预处理操作后的所述黄页需求相关信息,在黄页索引库 15' 中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。例如,判断装置 12' 获取的用户的黄页需求相关信息为“上海交通大学”,则预处理装置 16' 可以对此进行预处理操作,将其变换成“上海交通大学”(或“上海交通大学附属中学”)等;此后,查询装置 13' 再根据该预处理操作后的黄页需求相关信息“上海交通大学”(或“上海交通大学附属中学”),在黄页索引库 15' 中进行匹配查询,获取与之相匹配的一个或多个匹配网页,所述一个或多个匹配网页中包含有“上海交通大学”(或“上海交通大学附属中学”)的黄页信息。

[0052] 优选地,所述预处理操作包括但不限于以下至少一项:a) 近义词替换;b) 同义词替换;c) 词义拓展。近义词替换例如“学校”替换为“小学”、“中学”或“大学”等。在上例中,预处理装置 16' 可以依次将“上海交通大学”替换成“上海交通小学”、“上海交通中学”和“上海交通大学”,查询装置 13' 依次将替换后的黄页需求相关信息在黄页索引库 15' 中进行匹配查询。同义词替换例如将“上海交大”替换为“上海交通大学”等;词义拓展例如将“公司”拓展为“有限公司”、“有限责任公司”等。本领域技术人员应能理解上述预处理操作仅为举例,其他现有的或今后可能出现的预处理操作如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0053] 在另一个优选的实施例中(参照图 3),网络设备 1 还可以包括排序装置 17'。其中,第一获取装置 11'、判断装置 12'、查询装置 13' 和第一提供装置 14' 分别与图 2 所示对应装置 11-14 相同,故此处不再赘述,并通过引用的方式包含于此。

[0054] 其中,排序装置 17' 用于根据预设规则,确定所述一个或多个匹配网页的优先级;其中,第一提供装置 14' 根据所述优先级,将所述一个或多个匹配网页提供给所述用户设备。具体地,排序装置 17' 根据预设规则,例如根据网络设备 1 获取的或接收来自用户设备 2 统计的,或接收来自其他产品或第三方设备统计的所述匹配网页的累计点击量、所述匹配网页的搜索排名等,确定查询装置 13' 所获取的所述一个或多个匹配网页的优先级;此后,第一提供装置 14' 根据所述排序装置 17' 所确定的优先级,通过页面技术,例如 JSP、ASP、PHP 技术,将所述一个或多个匹配网页以链接、摘要或缩略图等形式,按一定顺序和格式,嵌入在搜索结果页面之中或之上,提供给所述用户设备 2。例如,排序装置 17' 可以根据所述一个或多个匹配网页的权威性,确定所述一个或多个匹配网页的优先级,如,对于同样包含“百度在线网络技术有限公司”的黄页信息的两个匹配网页“百度在线网络技术有限公司”和“百度在线网络技术有限公司_百度在线网络技术有限公司简介_艾瑞网行业公司”,由于前者的权威性更高,因此其优先级较后者高。第一提供装置 14' 根据排序装置 17' 确定的所述一个或多个匹配网页的优先级,例如按照优先级由大到小的顺序自上而下或由内而外,以链接、摘要或缩略图等形式,将排序后的所述一个或多个匹配网页提供给用户设备 2,供用户进行浏览。接前例,随后,第一提供装置 14' 根据这两个匹配网页的优先级,将这两个匹配网页提供给所述用户设备,其中,匹配网页“百度在线网络技术有限公司”相对“百度在线网络技术有限公司_百度在线网络技术有限公司简介_艾瑞网行业公司”更为优先展示。本领域技术人员应能理解上述将匹配网页提供给用户的方式仅为举例,其他现有的或今后可能出现的将匹配网页提供给用户的方式如可适用于本发明,也应包含在本发明保护

范围以内,并在此以引用方式包含于此。

[0055] 优选地,所述预设规则基于包括但不限于以下至少任一项:a)所述匹配网页的权威性;b)所述匹配网页的累计点击量;c)所述匹配网页的搜索排名;d)所述匹配网页中黄页信息的更新时间。排序装置 17'可以根据所述一个或多个匹配网页中各网页的权威性,确定所述一个或多个匹配网页的优先级。例如,对于同样包含“百度在线网络技术有限公司”的黄页信息的两个匹配网页“百度在线网络技术有限公司”和“百度在线网络技术有限公司_百度在线网络技术有限公司简介_艾瑞网行业公司”,由于前者的权威性更高,因此其优先级较后者高。排序装置 17'还可以根据网络设备 1 获取的或接收来自用户设备 2 统计的,或接收来自其他产品或第三方设备获取的所述匹配网页的累计点击量,确定所述一个或多个匹配网页的优先级。例如,对于查询装置 13'所获取的一个或多个匹配网页,排序装置 17'根据所述一个或多个匹配网页中各网页的累计点击量确定其优先级,累计点击量多的匹配网页优先级较高,累计点击量少的匹配网页则优先级较低。排序装置 17'还可以根据网络设备 1 获取的或接收来自用户设备 2 统计的,或接收来自其他产品或第三方设备获取的所述匹配网页的搜索排名,确定所述一个或多个匹配网页的优先级。例如,对于查询装置 13'所获取的一个或多个匹配网页,排序装置 17'根据用户以所述匹配网页作为搜索的搜索排名,确定其优先级,搜索排名靠前的匹配网页优先级较高,搜索排名靠后的匹配网页则优先级较低。排序装置 17'还可以根据所述匹配网页中黄页信息的更新时间,确定所述一个或多个匹配网页的优先级。例如,对于查询装置 13'所获取的一个或多个匹配网页,排序装置 17'根据所述匹配网页中黄页信息的更新时间,提高最近更新的匹配网页的优先级。又如,对于查询装置 13'匹配查询获取的一个或多个匹配网页,其上对于同一个机构的黄页信息可能不同,这时,可以根据网页上该黄页信息的更新时间进行排序,更新时间越近的优先级越高。优选地,排序装置 17'还可以根据上述所述匹配网页的权威性、所述匹配网页的累计点击量、所述匹配网页的搜索排名和所述匹配网页中黄页信息的更新时间的任意组合,通过例如分配给每一项因素一个权重,使用加权等方式,确定所述一个或多个匹配网页的优先级。本领域技术人员应能理解上述预设规则所基于的因素仅为举例,其他现有的或今后可能出现的预设规则所基于的因素如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0056] 在一个优选的实施例中(参照图 3),网络设备 1 还可以包括第四获取装置(未示出)。所述第四获取装置用于获取搜索引擎基于所述查询序列提供的第一搜索结果;其中,所述第一提供装置还用于将所述一个或多个匹配网页和所述第一搜索结果提供给所述用户设备 2,其中,优先提供所述一个或多个匹配网页。具体地,第一提供装置 14'除将包含有黄页信息的匹配网页提供给用户设备 2 之外,还可以将第四获取装置获取的搜索引擎基于所述查询序列提供的第一搜索结果提供给所述用户设备 2,其中,优先提供所述一个或多个匹配网页,例如,将所述一个或多个匹配网页以比所述第一搜索结果更高的优先级提供给所述用户设备 2。

[0057] 本领域技术人员应能理解,上述第一获取装置、第二获取装置、第三获取装置和第四获取装置仅为示例,在实践中,它们可以是四个独立的模块,或者任意两个装置集成在一个模块中,或者任意三个装置集成在一个模块中,也可全部集成在一个模块中。

[0058] 优选地,网络设备 1 还可以包括合并装置(未示出),合并装置用于将所述一个或

多个匹配网页与所述第一搜索结果进行合并去冗余处理,以获得第二搜索结果;此后,第一提供装置 14' 将所述第二搜索结果提供给所述用户设备 2。由于基于黄页需求相关信息匹配查询得到的匹配网页与基于查询序列搜索得到的第一搜索结果可能存在相同的网页,需要对此进行合并去冗余处理,以免提供给所述用户设备 2 相同的网页。例如,用户输入查询序列“百度公司”,网络设备 1 通过查询装置 13' 在索引库中匹配查询得到的匹配网页中包含有“百度在线网络技术有限公司”这个网页,而通过第四获取装置获取得到的搜索引擎基于查询序列“百度公司”提供的第一搜索结果中也同样包含有“百度在线网络技术有限公司”这个网页,合并装置可以将这两个网页进行合并去冗余处理,只留下其中一个匹配网页;此后,第一提供装置 14' 将合并处理后的匹配网页提供给所述用户设备 2。

[0059] 优选地,上述基于查询序列获取黄页信息的设备可以与现有的搜索引擎相结合,构成一种搜索引擎。搜索引擎(search engine)是指万维网环境中的信息检索系统,根据一定的策略、运用特定的计算装置搜集互联网上的信息,在对信息进行组织和处理后,将处理后的信息显示给用户。现有的搜索引擎可以是例如百度、Google、雅虎、有道、腾讯、搜狗等。

[0060] 图 4 示出根据本发明另一个方面的用于基于查询序列获取黄页信息的方法流程图。

[0061] 其中,在步骤 S1 中,网络设备 1 获取用户通过用户设备输入的查询序列。具体地,用户通过与用户设备 2 的交互方式,包括但不限于键盘、鼠标、遥控器、触摸板、或手写设备,在浏览器软件、应用程序或客户端软件中输入查询序列;在步骤 S1 中,网络设备 1 通过页面技术,例如 JSP、ASP、PHP 技术,实时地获取用户输入的查询序列。本领域技术人员应能理解上述页面技术仅为举例,其他现有的或今后可能出现的页面技术如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。本领域技术人员还应能理解上述获取用户浏览行为的技术仅为举例,其他现有的或今后可能出现的获取用户浏览行为的技术如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0062] 在步骤 S2 中,网络设备 1 根据所述查询序列,判断所述查询序列是否包含黄页需求相关信息。具体地,在步骤 S2 中,网络设备 1 根据在步骤 S1 中获取的用户输入的查询序列,通过例如词义分析、查询用户历史查询操作等方式,判断所述查询序列是否包含黄页需求相关信息。以词义分析为例,所述查询序列中可能包含多个黄页需求相关信息,当网络设备 1 在步骤 S2 中检测到其中一个黄页需求相关信息时,即可以判断所述查询序列包含黄页需求相关信息,但此后,网络设备 1 仍需继续执行操作,以获取所述查询序列中的全部黄页需求相关信息,以方便在步骤 S3 中根据所述全部黄页需求相关信息进行匹配查询。优选地,所述在步骤 S2 中,网络设备 1 也可以先进行获取的操作,获取所述查询序列中的全部黄页需求相关信息,再判断有无获取到所述黄页需求相关信息,若有,则网络设备 1 执行步骤 S3。例如,网络设备 1 在步骤 S1 中获取到用户输入的查询序列为“百度地址”,在步骤 S2 中,检测到该查询序列中包含的“百度”为黄页需求相关信息,判断该查询序列包含黄页需求相关信息,此后,网络设备 1 继续进行检测,获取该查询序列中的另一个黄页需求相关信息“地址”。又如,在步骤 S2 中,网络设备 1 可以获取该查询序列中的全部黄页需求相关信息“百度”和“地址”,再判断该查询序列中包含有黄页需求相关信息。本领域技术人员应能

理解上述判断查询序列是否包含黄页需求相关信息的方式仅为举例,其他现有的或今后可能出现的判断查询序列是否包含黄页需求相关信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0063] 在步骤 S3 中,当所述查询序列包含黄页需求相关信息时,网络设备 1 根据所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。具体地,当网络设备 1 在步骤 S2 中判断得出所述查询序列包含黄页需求相关信息时,在步骤 S3 中,网络设备 1 根据所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。在此,所述黄页索引库中存储有大量包含有黄页信息网页,例如,网络设备 1 可以根据大量的网页和其中的黄页信息,通过建立或更新数据表等形式,建立或更新所述黄页索引库,在所述数据表中,可以是一个黄页信息对应一个网页,也可以是一个黄页信息对应包含有该黄页信息的几个网页。在步骤 S3 中,网络设备 1 根据所述黄页需求相关信息,在黄页索引库中进行匹配查询,首先获取与所述黄页需求相关信息相匹配的黄页信息,再根据所述黄页信息,匹配获取一个或多个与所述黄页信息相对应的匹配网页。例如,在黄页索引库中,存储有“百度时代网络技术(北京)有限公司”、“百度在线网络技术有限公司”、“百度时代网络技术(北京)有限公司客户发展部”三个标题的网页及其中的黄页信息,用户输入查询序列“百度北京”,在步骤 S3 中,网络设备 1 根据其中的黄页需求相关信息“百度北京”,在黄页索引库中匹配到与之相对应的上述三个网页的黄页信息,并根据所述黄页信息,匹配获取到上述三个网页。本领域技术人员应能理解上述获取匹配网页的方式仅为举例,其他现有的或今后可能出现的获取匹配网页的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0064] 在步骤 S4 中,网络设备 1 将所述一个或多个匹配网页提供给所述用户设备 2。具体地,在步骤 S4 中,网络设备 1 根据在步骤 S3 中获取的一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配,通过页面技术,例如 JSP、ASP、PHP 技术,将所述一个或多个匹配网页以链接、摘要或缩略图等形式,按一定顺序和格式,嵌入在搜索结果页面中,提供给所述用户设备 2,供用户进行浏览。接前例,在步骤 S4 中,网络设备 1 将在步骤 S3 中获取的“百度时代网络技术(北京)有限公司”、“百度在线网络技术有限公司”、“百度时代网络技术(北京)有限公司客户发展部”三个标题的网页,以链接、摘要或缩略图等形式,按一定顺序和格式,嵌入于搜索结果页面之中或之上,提供给所述用户设备 2。本领域技术人员应能理解上述将匹配网页提供给用户设备的方式仅为举例,其他现有的或今后可能出现的将推荐信息提供给用户设备的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0065] 优选地,上述各个步骤之间是持续不断工作的。具体地,网络设备 1 在步骤 S1 中获取用户通过用户设备 2 输入的查询序列;随后,在步骤 S2 中,网络设备 1 根据所述查询序列,判断所述查询序列是否包含黄页需求相关信息;当所述查询序列包含黄页需求相关信息时,网络设备 1 在步骤 S3 中根据所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配;接着,在步骤 S4 中,网络设备 1 再将所述一个或多个匹配网页提供给所

述用户设备。在此,本领域技术人员应理解“持续”是指各步骤分别按照设定的或实时调整的工作模式要求进行查询序列的获取、黄页需求相关信息的判断、匹配网页的获取及提供,直至该用户在较长时间内停止通过用户设备 2 输入查询序列。

[0066] 优选地,在步骤 S7(未示出)中,网络设备 1 还可以将所述黄页信息以突显的形式提供给所述用户设备。具体地,在步骤 S7 中,网络设备 1 根据在步骤 S3 中获取的一个或多个匹配网页中的黄页信息,通过页面技术,例如 JSP、ASP、PHP 技术,将所述黄页信息以摘要、悬浮框或缩略图等形式,按一定顺序和格式,嵌入于搜索结果页面之中或之上,提供给所述用户设备 2,供用户进行浏览。在此,黄页信息的展现应与前述匹配网页的展现相对应,例如当所述黄页信息以摘要的形式展现在搜索结果页面时,可以在所述匹配网页的链接的下方展现给用户。优选地,还可以将所述黄页信息加粗、加黑,或改变字体、字形、颜色,或增加背景颜色等方式,将其进行突显展现。当以悬浮框的形式提供给用户设备时,可以是在页面的某个固定的区域显示,也可以由用户设定或根据浏览对象的位置确定,其窗口大小可以进行调节。例如,将所述黄页信息以悬浮框的形式显示在用户设备 2 的输入框中,其可以是透明的以不影响用户的输入,并随着用户的输入操作不断变换其所处的位置,例如随着输入序列变长,所述悬浮框不断右移。优选地,所述悬浮框还可以根据光标与之的距离自动调整自身的透明度,以不影响用户的正常浏览;当光标移动到距所述悬浮框的设定距离范围外时隐藏所述悬浮框;当在所述悬浮框之外点击时,隐藏所述悬浮框。本领域技术人员应能理解上述提供黄页信息的方式仅为举例,其他现有的或今后可能出现的提供黄页信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0067] 优选地,在步骤 S8(未示出)中,网络设备 1 还可以对所述黄页信息进行规范化处理,以获得规范化的黄页信息;随后,在步骤 S7(未示出)中,网络设备 1 将所述规范化的黄页信息以突显的形式提供给所述用户设备 2。在步骤 S3 中,网络设备 1 匹配查询到的匹配网页中的黄页信息可能存在多种问题:1)形式上并不统一,例如有的网页在电话号码前加上了区号,有的没有加区号,再如有的匹配网页中的黄页信息包含电子邮箱地址,有的则不包含;2)表述上不规范,如将“中国人民大学附属中学”记为“人大附中”。在步骤 S8 中,网络设备 1 可以将所述黄页信息进行规范化处理,以获得形式统一、表述规范的黄页信息。例如,在步骤 S8 中,网络设备 1 根据预置的黄页模板,对所述黄页信息进行规范化处理,例如将电话号码统一按照“国际区号-国内区号-电话号码”的形式显示,又如假设规定提供给用户的黄页信息包含有项目“机构名称、地址、邮编、网址、电话、传真、法定代表人”并按照一定的格式排版,若该网页中不包含某一项的信息,则相应的项目空置。又如,在步骤 S8 中,网络设备 1 可以根据对应于同一个黄页需求相关信息的多个匹配网页中的黄页信息,通过归类去重处理得到对应于该黄页需求相关信息的黄页模板,然后根据此黄页模板对这多个匹配黄页中的行黄页信息进行规范化,如将这些黄页信息按归类去重处理所得到的黄页模板进行格式排版,将该黄页信息中的信息项填入该黄页模板的对应项目中,若该黄页信息不包含某个信息项,则空置该黄页模板中的对应项目。再如,在步骤 S8 中,网络设备 1 还可以通过自然语言处理技术,使得黄页信息中的信息项表述更为规范,例如在步骤 S8 中,网络设备 1 利用自然语言处理技术,将机构名称按 Term 切分,诸如“人大附中”,可分割为“人大”、“附中”;并结合同义词列表、别名列表,可将其规范化为“中国人民大学附属中学”。此后,在步骤 S7 中,网络设备 1 再通过页面技术,例如 JSP、ASP、PHP 技术,将所述模板

嵌入于返回给用户设备的搜索结果页面之中或之上,以突显的形式提供给用户设备 2。本领域技术人员应能理解,在步骤 S8 中,网络设备 1 还可以通过黄页模板与自然语言处理技术的结合,获取形式统一且表述规范的黄页信息。本领域技术人员还应能理解上述规范化处理黄页信息的方式仅为举例,其他现有的或今后可能出现的规范化处理黄页信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0068] 优选地,在步骤 S9(未示出)中,网络设备 1 还可以根据所述黄页需求相关信息,并结合所述查询序列,获取所述查询序列中的黄页附加需求相关信息;根据所述黄页信息与所述黄页附加需求相关信息,获取与所述黄页附加需求相关信息相匹配的黄页附加信息;随后,在步骤 S7(未示出)中,网络设备 1 将所述黄页信息与所述黄页附加信息以突显的形式提供给所述用户设备 2。在用户输入的查询序列中,除与机构名称有关、与机构联系方式有关或与机构法人有关的黄页需求相关信息外,还可能包含黄页附加需求相关信息,例如与该机构有关的地图相关信息、路线相关信息等需求。具体地,在步骤 S9 中,网络设备 1 根据在步骤 S2 中所获取的黄页需求相关信息,并结合在步骤 S1 中所获取的查询序列,通过词义分析、数据库查询等方式,获取所述查询序列中的黄页附加需求相关信息;此后,网络设备 1 再根据所述黄页信息与所述黄页附加需求相关信息,通过数据库查询或利用第三方设备提供的应用程序接口(API)从第三方设备获取等方式,获取与所述黄页附加需求相关信息相匹配的黄页附加信息;随后,在步骤 S7 中,网络设备 1 通过页面技术,例如 JSP、ASP、PHP 技术,将所述黄页信息与所述黄页附加信息嵌入于返回给用户设备的搜索结果页面之中或之上,以突显的形式提供给用户设备 2。例如,用户通过用户设备 2 输入查询序列“百度北京怎么去”,在步骤 S2 中,网络设备 1 通过词义分析,获取到该用户的黄页需求相关信息为“百度北京”,在步骤 S9 中,网络设备 1 根据该用户的黄页需求相关信息,并结合该用户输入的查询序列,通过词义分析,获取到黄页附加需求相关信息为“地图”(或“乘车路线”);此后,网络设备 1 再根据“百度北京”的黄页信息与所述黄页附加需求相关信息,通过调用第三方地图搜索设备获取“百度北京”的地图信息(或者通过数据库查询获得“百度北京”周边的乘车信息,如轨道交通与公交线路);随后,在步骤 S7 中,网络设备 1 再将所述“百度北京”的黄页信息与地图信息(或乘车路线)提供给用户设备 2。优选地,在步骤 S9 中,网络设备 1 还可以取用户输入的查询序列所包括的另一位置信息或者用户当前的定位信息(假设用户设备 2 包含有 GPS 等地理信息模块等)作为一个端点,提供该端点与所述黄页信息所对应的另一个端点之间的详细的路线图。本领域技术人员应能理解上述获取与提供黄页附加信息的方式仅为举例,其他现有的或今后可能出现的获取或提供黄页附加信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0069] 优选地,在步骤 S10(未示出)中,网络设备 1 还可以根据所述黄页需求相关信息,并结合所述查询序列,获取所述查询序列中的黄页焦点需求相关信息;根据所述黄页信息与所述黄页焦点需求相关信息,获取与所述黄页焦点需求相关信息相匹配的黄页焦点信息;此后,在步骤 S7 中,网络设备 1 将所述黄页焦点信息以突显的形式提供给所述用户设备。由于对于某一部分用户来讲,其可能只需要黄页信息中的黄页焦点信息,例如地址、电话、网站中的一个或多个,网络设备 1 需要获取该用户的黄页焦点需求相关信息,并进一步获取与之相匹配的黄页焦点信息。具体地,在步骤 S10 中,网络设备 1 根据在步骤 S2 中所获取的黄页需求相关信息,并结合在步骤 S1 中所获取的查询序列,获取所述查询序列中的

黄页焦点需求相关信息;此后,在步骤 S10 中,网络设备 1 根据所述黄页信息与所述黄页焦点需求相关信息,通过词义分析、数据库查询等方式,获取与所述黄页焦点需求相关信息相匹配的黄页焦点信息;随后,在步骤 S7 中,网络设备 1 通过页面技术,例如 JSP、ASP、PHP 技术,将所述黄页焦点信息嵌入于返回给用户设备 2 的搜索结果页面之中或之上,以突显的形式提供给用户设备 2。例如,用户通过用户设备 2 输入查询序列“联想客服电话”,在步骤 S2 中,网络设备 1 通过词义分析,获取到该用户的黄页需求相关信息为“联想 lenovo”,在步骤 S10 中,网络设备 1 根据该用户的黄页需求相关信息,并结合该用户输入的查询序列,通过词义分析,获取到黄页焦点需求相关信息为“客服电话”;此后,在步骤 S10 中,网络设备 1 再通过词义分析、数据库查询等方式,获取与所述黄页焦点需求相关信息相匹配的黄页焦点信息“8008108888”;随后,在步骤 S7 中,网络设备 1 再通过页面技术,例如 JSP、ASP、PHP 技术,将所述黄页焦点信息“8008108888”嵌入于返回给用户设备 2 的搜索结果页面之中或之上,以突显的形式提供给用户设备 2。例如,将所述黄页焦点信息“8008108888”以悬浮框的形式显示在用户设备 2 的输入框中,其可以是透明的以不影响用户的输入,并随着用户的输入操作不断变换其所处的位置,例如随着输入序列变长,所述悬浮框不断右移。本领域技术人员应能理解上述获取与提供黄页焦点信息的方式仅为举例,其他现有的或今后可能出现的获取或提供黄页焦点信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0070] 优选地,在步骤 S7(未示出)中,网络设备 1 将黄页信息提供给用户设备 2 的突显方式,包括但不限于,以下至少一项:摘要、悬浮框和缩略图等形式。在步骤 S7 中,网络设备 1 可以通过页面技术,例如 JSP、ASP、PHP 技术,将在步骤 S3 中获取的黄页信息以摘要的形式提供给所述用户设备 2,黄页信息的展现应与该黄页信息所属的匹配网页的展现相对应,可以在所述匹配网页的链接的下方展现给用户。优选地,还可以将所述黄页信息加粗、加黑,或改变字体、字形、颜色,或增加背景颜色等方式,将其进行突显展示。在步骤 S7 中,网络设备 1 还可以将在步骤 S3 中获取的黄页信息,以悬浮框的形式提供给所述用户设备 2,用户通过点击网页上特定的按钮或用户自定义的按钮或按键,或者将光标悬停指向搜索结果页面中的链接时,打开所述悬浮框,以浏览所述黄页信息。优选地,所述悬浮框可以是在页面的某个固定的区域显示,也可以由用户设定或根据浏览对象的位置确定,其窗口大小可以进行调节。所述悬浮框还可以根据光标与之的距离自动调整自身的透明度,以不影响用户的正常浏览;当光标移动到距所述悬浮框的设定距离范围外时隐藏所述悬浮框;当在所述悬浮框之外点击时,隐藏所述悬浮框。例如,将所述黄页信息以悬浮框的形式显示在用户设备 2 的输入框中,其可以是透明的以不影响用户的输入,并随着用户的输入操作不断变换其所处的位置,例如随着输入序列变长,所述悬浮框不断右移。在步骤 S7 中,网络设备 1 还可以将在步骤 S3 中获取的黄页信息,以缩略图的形式提供给所述用户设备 2,用户可以通过点击特定的或自定义的按钮放大或缩小所述缩略图,以浏览所述黄页信息。在此,本领域技术人员应能理解,上述突显的形式同样适用于网络设备 1 在步骤 S7 中将规范化的黄页信息、黄页信息与黄页附加信息以突显的形式提供给所述用户设备 2 本领域技术人员还应能理解上述突显的形式仅为举例,其他现有的或今后可能出现的突显的形式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0071] 优选地,在步骤 S2 中,网络设备 1 还可以对所述查询序列进行词义分析,当确定所

述查询序列符合预定条件时,判断所述查询序列包含黄页需求相关信息,其中,所述预定条件包括但不限于以下至少一项:a)所述查询序列包含与机构名称有关的特定类别关键词;b)所述查询序列包含与机构联系方式有关的特定类别关键词;c)所述查询序列包含与机构法人有关的特定类别关键词。具体地,在步骤 S2 中,网络设备 1 根据在步骤 S1 中获取的用户输入的查询序列,通过词义分析,判断所述查询序列中是否包含与机构名称有关的特定类别关键词,例如机构的部分名称或全部名称,是否包含与机构联系方式有关的特定类别关键词,例如电话、邮箱、传真、网址等,是否包含与机构法人有关的特定类别关键词,例如法人代表、法定代表人等;当所述查询序列包含上述特定类别关键词时,判断所述查询序列包含黄页需求相关信息。本领域技术人员应能理解上述判断查询序列是否包含黄页需求相关信息的预定条件仅为举例,其他现有的或今后可能出现的判断查询序列是否包含黄页需求相关信息的预定条件如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0072] 更优选地,在步骤 S2 中,网络设备 1 还可以根据所述查询序列,并结合该用户的历史查询操作,判断所述查询序列是否包含黄页需求相关信息。具体的,在步骤 S2 中,网络设备 1 根据在步骤 S1 中获取的用户输入的查询序列,并结合该用户的历史查询操作,如用户曾使用的查询关键词、用户对包含黄页信息的网页的点击浏览记录,判断所述查询序列是否包含黄页需求相关信息。例如,在步骤 S2 中,网络设备 1 根据用户当前的查询序列为“百度”,结合该用户的历史查询操作,如曾经点击包含有“百度公司”黄页信息的网页的 URL,则判断该查询序列包括黄页需求相关信息。又如,在步骤 S2 中,网络设备 1 根据用户当前的查询序列为“百度”,结合该用户的历史查询操作,如该用户曾经变换查询序列,例如用户输入查询序列“百度”,在获得的搜索结果页面中点击联想查询序列“百度公司的地址”,后者为明确的黄页需求,则判断该查询序列包括黄页需求相关信息。在步骤 S2 中,网络设备 1 还可以对所述查询序列进行词义分析,并结合该用户的历史查询操作,当确定该查询序列满足预定条件时,判断该查询序列包含黄页需求相关信息。例如,在步骤 S2 中,网络设备 1 根据用户当前的查询序列为“百度”,结合该用户的历史查询操作,如该用户曾经变换查询序列,例如用户输入查询序列“百度”,在获得的搜索结果页面中点击联想查询序列“百度公司法人代表”,后者为明确的黄页需求,则判断该查询序列包括黄页需求相关信息。本领域技术人员还应能理解上述判断查询序列是否包含黄页需求相关信息的方式仅为举例,其他现有的或今后可能出现的判断查询序列是否包含黄页需求相关信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。在此,该用户的所述历史查询操作既可以是网络设备 1 独立获取的,也可以是接受来自其他产品或第三方设备所完成的对该用户的所述历史查询操作的获取。例如,可通过检测该用户使用网络设备 1 的搜索服务过程中的操作,或通过向用户设备 2 询问并获取该用户的 COOKIE 信息来获得该用户的历史查询操作,优选地,可赋予该用户一个临时 ID 并为其建立一个历史查询操作库,将此次搜索过程中用户的历史记录保存于该临时历史查询操作库中。本领域技术人员还应能理解上述获取用户的历史查询操作的方式仅为举例,其他现有的或今后可能出现的获取用户的历史查询操作的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0073] 优选地,在步骤 S11(未示出)中,网络设备 1 还可以根据大量网页及其中的黄页

信息,建立或更新所述黄页索引库。具体地,在步骤 S11 中,网络设备 1 对通过爬虫工具等网页抓取工具直接抓取的大量网页,或者从第三方设备中获取的存储于该第三方设备的大量网页,进行特征分析,以获取包含黄页信息的潜在目标页,并根据这些潜在目标页及其中的黄页信息,建立或更新所述黄页索引库。在此,网络设备 1 在步骤 S11 中确定潜在目标页的方式,包括但不限于:1) 根据网址中包含有例如“show_company”、“huangye”、“yellowpage”等表意信息的网页为潜在目标页;2) 根据网页 Title 和 Meta 信息域中含有“公司、企业、公司介绍、企业简介”等表意信息的网页为潜在目标页;3) 根据在网页原文中提取有“地址、电话、传真、邮编、联系人、名称”等表意的信息对的网页为潜在目标页;4) 或者根据其它的指向信息,例如 anchor(锚文本)信息中包含有以上表意信息等的网页为潜在目标页;5) 上述几种方式的任意组合。网络设备 1 在步骤 S11 中通过相互补充验证,提取潜在目标页中的有效机构名称、地址信息、邮编、电话、联系人、传真等黄页信息。将所述黄页信息与所述潜在目标页通过建立或更新数据表等形式,建立或更新所述黄页索引库。在所述数据表中,可以是一个黄页信息对应一个网页,也可以是一个黄页信息对应包含有该黄页信息的几个网页。例如,在黄页索引库中,“百度在线网络技术有限公司”的黄页信息“公司全称:百度在线网络技术有限公司;网址:http://www.baidu.com;电话:86-10-82621188;电子邮件:webmaster@baidu.com”可以只对应“百度在线网络技术有限公司”一个网页,也可以同时对应“百度在线网络技术有限公司百度在线网络技术有限公司简介艾瑞网行业公司”这个网页。本领域技术人员应能理解上述建立或更新黄页索引库的方式仅为举例,其他现有的或今后可能出现的建立或更新黄页索引库的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0074] 优选地,在步骤 S3 中,当所述查询序列包含黄页需求相关信息时,网络设备 1 还可以根据所述黄页需求相关信息,并结合该用户的用户相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。具体地,当网络设备 1 在步骤 S2 中判断所述查询序列包含黄页需求相关信息时,在步骤 S3 中根据所述黄页需求相关信息,并结合该用户的用户相关信息,如该用户的搜索历史记录或用户设备的地址,在黄页索引库中进行匹配查询,以获得与用户实际需求相匹配的一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。例如,网络设备 1 在步骤 S3 中先将用户相关信息的关键词与黄页需求相关信息组合成第二查询序列,随后,网络设备 1 再根据所述第二查询序列,在黄页索引库中进行查询匹配,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。在此,所述用户相关信息既可以是网络设备 1 独立获取的,也可以是接受来自用户设备 2 记录的,或者接受来自其他产品或第三方设备所获取的用户相关信息。本领域技术人员应能理解上述获取用户相关信息的方法仅为举例,其他现有的或今后可能出现的获取用户相关信息的方法如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0075] 优选地,所述用户相关信息包括但不限于以下至少一项:a) 用户属性;b) 用户偏好;c) 用户搜索历史记录;d) 用户设备的地址。用户属性包括但不限于,用户预先输入的个人背景信息,例如性别、年龄、职业、宗教、国籍、教育程度等。当所述查询序列包含黄页需求相关信息时,网络设备 1 在步骤 S3 中根据所述黄页需求相关信息,并结合该用户的用

户属性,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。例如,用户输入查询序列“新东方”进行查询,当用户为学生时,提供有关“新东方学校”的黄页信息及相应的匹配网页,当用户为厨师时,提供有关“新东方厨师学校”的黄页信息及相应的匹配网页。用户偏好包括但不限于,用户预先设置的一些偏好选择,例如关注的机构类型、行业信息等。当所述查询序列包含黄页需求相关信息时,网络设备 1 在步骤 S3 中根据所述黄页需求相关信息,并结合该用户的用户偏好,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。例如,用户设置其偏好的机构类型为客服公司,则当用户输入查询序列“百度北京”时,优先提供有关“百度公司客服”的黄页信息及相应的网页,当用户设置其偏好的机构类型为研发公司时,其输入查询序列“百度北京”,优先提供有关“百度在线网络技术(北京)有限公司”等的黄页信息及相应的网页。用户搜索历史记录,其由系统来记录和统计用户先前的搜索和点击浏览历史,从而可判断或预测用户的需求。其中,当用户为登录用户时,网络设备 1 可通过该用户的 ID 在网络端保存有该用户相关信息的用户特征库中进行查询;当用户为非登录用户时,可通过检测该用户使用本网络设备 1 的搜索服务过程中的动作,或通过向用户设备 2 询问并获取该用户的 COOKIE 信息来获得,优选地,可赋予该用户一个临时 ID 并为其建立一个用户特征库,将此次搜索过程中用户的历史记录保存于该临时用户特征库中。例如,用户输入查询序列“复旦大学”,提供的匹配网页可能包括关于“复旦大学”、“复旦大学医学院”、“复旦大学出版社”等的黄页信息,当用户点击“复旦大学医学院”时,网络设备 1 或用户设备 2 记录下用户的此次浏览操作记录,当用户再次输入“复旦大学”或其他大学的名称作为查询序列时,网络设备 1 优先将关于“某某大学出版社”的黄页信息及相应的匹配网页提供给该用户。用户设备的地址,其可以反映该用户当前所处地域,可由网络设备 1 实时检测用户设备 2 的 IP 地址来获取。例如,用户通过用户设备 2 输入查询序列“新东方语言学校”,网络设备 1 检测到该用户设备 2 的 IP 地址位于北京,则优先将北京新东方语言学校的黄页信息及相应的匹配网页提供给该用户。本领域技术人员应能理解上述用户相关信息仅为举例,其他现有的或今后可能出现的用户相关信息如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0076] 优选地,在步骤 S12(未示出)中,网络设备 1 还可以获取该用户对所述一个或多个匹配网页的浏览操作;根据所述浏览操作,更新该用户的所述用户相关信息。具体地,在步骤 S12 中网络设备 1 根据与用户设备 2 及用户的交互,例如通过 JSP、ASP、PHP 等页面技术,获取该用户对所述一个或多个匹配网页的浏览操作;并根据这些浏览操作更新该用户的用户相关信息。在此,网络设备 1 在步骤 S12 中更新用户相关信息的方式包括但不限于:1) 网络设备 1 根据用户更改的预设信息,更新该用户的用户相关信息;2) 网络设备 1 可在用户一次浏览操作结束后记录用户在此浏览过程中的所有动作,例如点击、浏览或关闭;3) 网络设备 1 在用户一次浏览过程中记录用户每次动作并实时对网络端该用户的特征库进行更新,从而可在用户在此次浏览过程中点击其他黄页信息时提供最新的用户相关信息。例如,用户输入“复旦大学”,提供的匹配网页可能包括关于“复旦大学”、“复旦大学医学院”、“复旦大学出版社”等的黄页信息,当用户点击“复旦大学医学院”时,记录该用户的浏览操作并更新该用户的历史记录。本领域技术人员应能理解上述更新用户相关信息的方

式仅为举例,其他现有的或今后可能出现的更新用户相关信息的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0077] 图 5 示出根据本发明一个优选实施例的用于基于查询序列获取黄页信息的方法流程图。该方法还包括步骤 S5',其中,步骤 S1'-S4'分别与图 4 所示对应步骤 S1-S4 相同,故此处不再赘述,并通过引用的方式包含于此。

[0078] 其中,在步骤 S5'中,网络设备 1 根据所述黄页需求相关信息,对所述黄页需求相关信息进行预处理操作;其中,在步骤 S2'中,根据预处理操作后的所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。具体地,在步骤 S5'中,网络设备 1 根据在步骤 S2'中所获取的黄页需求相关信息,通过自然语言处理技术,例如近义词替换、同义词替换、词义拓展等方式,对所述黄页需求相关信息进行预处理操作;随后,在步骤 S3'中根据预处理操作后的所述黄页需求相关信息,在黄页索引库中进行匹配查询,获取一个或多个匹配网页,其中,所述一个或多个匹配网页中的黄页信息与所述黄页需求相关信息相匹配。例如,在步骤 S2'中,网络设备 1 获取的用户的黄页需求相关信息为“上海交通大学”,则在步骤 S5'中,网络设备 1 可以对此进行预处理操作,将其变换成“上海交通大学”(或“上海交通大学附属中学”)等;此后,在步骤 S3'中,网络设备 1 再根据该预处理操作后的黄页需求相关信息“上海交通大学”(或“上海交通大学附属中学”),在黄页索引库中进行匹配查询,获取与之相匹配的一个或多个匹配网页,所述一个或多个匹配网页中包含有“上海交通大学”(或“上海交通大学附属中学”)的黄页信息。

[0079] 优选地,所述预处理操作包括但不限于以下至少一项:a) 近义词替换;b) 同义词替换;c) 词义拓展。近义词替换例如“学校”替换为“小学”、“中学”或“大学”等。在上例中,网络设备 1 在步骤 S5'中可以依次将“上海交通大学”替换成“上海交通小学”、“上海交通中学”和“上海交通大学”,在步骤 S3'中依次将替换后的黄页需求相关信息在黄页索引库中进行匹配查询。同义词替换例如将“上海交大”替换为“上海交通大学”等;词义拓展例如将“公司”拓展为“有限公司”、“有限责任公司”等。本领域技术人员应能理解上述预处理操作仅为举例,其他现有的或今后可能出现的预处理操作如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0080] 在另一个优选的实施例中(参照图 5),该方法还可以包括步骤 S6'。其中,步骤 S1'-S4'分别与图 4 所示对应步骤 S1-S4 相同,故此处不再赘述,并通过引用的方式包含于此。

[0081] 其中,在步骤 S6'中,网络设备 1 根据预设规则,确定所述一个或多个匹配网页的优先级;随后,在步骤 S4'中,网络设备 1 根据所述优先级,将所述一个或多个匹配网页提供给所述用户设备。具体地,在步骤 S6'中,网络设备 1 根据预设规则,例如根据网络设备 1 获取的或接收来自用户设备 2 统计的,或接收来自其他产品或第三方设备统计的所述匹配网页的累计点击量、所述匹配网页的搜索排名等,确定在步骤 S3'中所获取的所述一个或多个匹配网页的优先级;随后,在步骤 S4'中,网络设备 1 根据在步骤 S6'中所确定的优先级,通过页面技术,例如 JSP、ASP、PHP 技术,将所述一个或多个匹配网页以链接、摘要或缩略图等形式,按一定顺序和格式,嵌入在搜索结果页面之中或之上,提供给所述用户设备 2。例如,在步骤 S6'中,网络设备 1 可以根据所述一个或多个匹配网页的权威性,确定所述一个或多

个匹配网页的优先级,如,对于同样包含“百度在线网络技术有限公司”的黄页信息的两个匹配网页“百度在线网络技术有限公司”和“百度在线网络技术有限公司_百度在线网络技术有限公司简介_艾瑞网行业公司”,由于前者的权威性更高,因此其优先级较后者高。在步骤 S4' 中,网络设备 1 根据在步骤 S6' 中确定的所述一个或多个匹配网页的优先级,例如按照优先级由大到小的顺序自上而下或由内而外,以链接、摘要或缩略图等形式,将排序后的所述一个或多个匹配网页提供给用户设备 2,供用户进行浏览。接前例,随后,在步骤 S4' 中,网络设备 1 根据这两个匹配网页的优先级,将这两个匹配网页提供给所述用户设备,其中,匹配网页“百度在线网络技术有限公司”相对“百度在线网络技术有限公司_百度在线网络技术有限公司简介_艾瑞网行业公司”更为优先展示。本领域技术人员应能理解上述将匹配网页提供给用户的方式仅为举例,其他现有的或今后可能出现的将匹配网页提供给用户的方式如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0082] 优选地,所述预设规则基于包括但不限于以下至少任一项:a) 所述匹配网页的权威性;b) 所述匹配网页的累计点击量;c) 所述匹配网页的搜索排名;d) 所述匹配网页中黄页信息的更新时间。在步骤 S6' 中,网络设备 1 可以根据所述一个或多个匹配网页中各网页的权威性,确定所述一个或多个匹配网页的优先级。例如,对于同样包含“百度在线网络技术有限公司”的黄页信息的两个匹配网页“百度在线网络技术有限公司”和“百度在线网络技术有限公司_百度在线网络技术有限公司简介_艾瑞网行业公司”,由于前者的权威性更高,因此其优先级较后者高。在步骤 S6' 中,网络设备 1 还可以根据网络设备 1 获取的或接收来自用户设备 2 统计的,或接收来自其他产品或第三方设备获取的所述匹配网页的累计点击量,确定所述一个或多个匹配网页的优先级。例如,对于网络设备 1 在步骤 S3' 中所获取的一个或多个匹配网页,在步骤 S6' 中根据所述一个或多个匹配网页中各网页的累计点击量确定其优先级,累计点击量多的匹配网页优先级较高,累计点击量少的匹配网页则优先级较低。网络设备 1 在步骤 S6' 中还可以根据网络设备 1 获取的或接收来自用户设备 2 统计的,或接收来自其他产品或第三方设备获取的所述匹配网页的搜索排名,确定所述一个或多个匹配网页的优先级。例如,对于网络设备 1 在步骤 S3' 中所获取的一个或多个匹配网页,在步骤 S6' 中根据用户以所述匹配网页作为搜索的搜索排名,确定其优先级,搜索排名靠前的匹配网页优先级较高,搜索排名靠后的匹配网页则优先级较低。网络设备 1 在步骤 S6' 中还可以根据所述匹配网页中黄页信息的更新时间,确定所述一个或多个匹配网页的优先级。例如,对于网络设备 1 在步骤 S3' 中所获取的一个或多个匹配网页,在步骤 S6' 中根据所述匹配网页中黄页信息的更新时间,提高最近更新的匹配网页的优先级。又如,对于网络设备 1 在步骤 S3' 中匹配查询获取的一个或多个匹配网页,其上对于同一个机构的黄页信息可能不同,这时,可以根据网页上该黄页信息的更新时间进行排序,更新时间越近的优先级越高。优选地,在步骤 S6' 中,网络设备 1 还可以根据上述所述匹配网页的权威性、所述匹配网页的累计点击量、所述匹配网页的搜索排名和所述匹配网页中黄页信息的更新时间的任意组合,通过例如分配给每一项因素一个权重,使用加权等方式,确定所述一个或多个匹配网页的优先级。本领域技术人员应能理解上述预设规则所基于的因素仅为举例,其他现有的或今后可能出现的预设规则所基于的因素如可适用于本发明,也应包含在本发明保护范围以内,并在此以引用方式包含于此。

[0083] 在一个优选的实施例中（参照图 5），在步骤 S13'（未示出）中，网络设备 1 还可以获取搜索引擎基于所述查询序列提供的第一搜索结果；随后，在步骤 S4' 中，网络设备 1 将所述一个或多个匹配网页和所述第一搜索结果提供给所述用户设备 2，其中，优先提供所述一个或多个匹配网页。具体地，在步骤 S4' 中，网络设备 1 除将包含有黄页信息的匹配网页提供给用户设备 2 之外，还可以将在步骤 S13' 中获取的搜索引擎基于所述查询序列提供的第一搜索结果提供给所述用户设备 2，其中，优先提供所述一个或多个匹配网页，例如，将所述一个或多个匹配网页以比所述第一搜索结果更高的优先级提供给所述用户设备 2。

[0084] 优选地，在步骤 S14'（未示出）中，网络设备 1 还可以将所述一个或多个匹配网页与所述第一搜索结果进行合并去冗余处理，以获得第二搜索结果；此后，在步骤 S4' 中，网络设备 1 将所述第二搜索结果提供给所述用户设备 2。由于基于黄页需求相关信息匹配查询得到的匹配网页与基于查询序列搜索得到的第一搜索结果可能存在相同的网页，需要对此进行合并去冗余处理，以免提供给所述用户设备 2 相同的网页。例如，用户输入查询序列“百度公司”，网络设备 1 通过在步骤 S3' 中在索引库中匹配查询得到的匹配网页中包含有“百度在线网络技术有限公司”这个网页，而通过步骤 S13'（未示出）获取到的搜索引擎基于查询序列“百度公司”提供的第一搜索结果中也同样包含有“百度在线网络技术有限公司”这个网页，在步骤 S14' 中，网络设备 1 可以将这两个网页进行合并去冗余处理，只留下其中一个匹配网页；此后，在步骤 S4' 中，网络设备 1 将合并处理后的匹配网页提供给所述用户设备 2。

[0085] 对于本领域技术人员而言，显然本发明不限于上述示范性实施细节，而且在不背离本发明的精神或基本特征的情况下，能够以其他的具体形式实现本发明。因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本发明的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。此外，显然“包括”一词不排除其他单元或步骤，单数不排除复数。装置权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一，第二等词语用来表示名称，而并不表示任何特定的顺序。

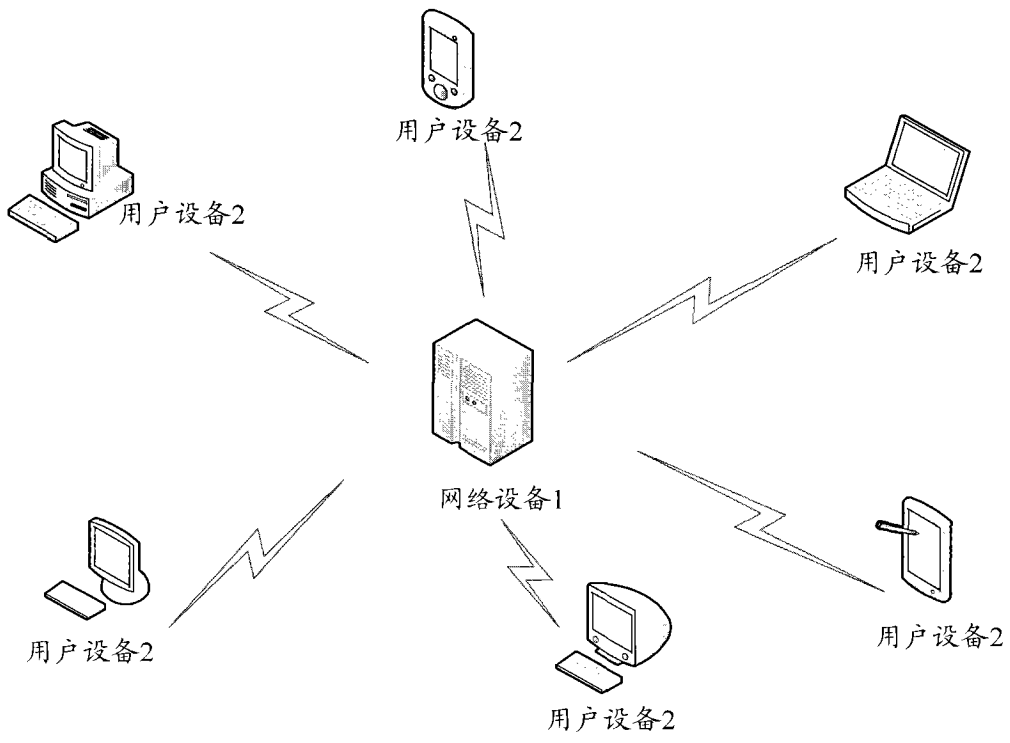


图 1

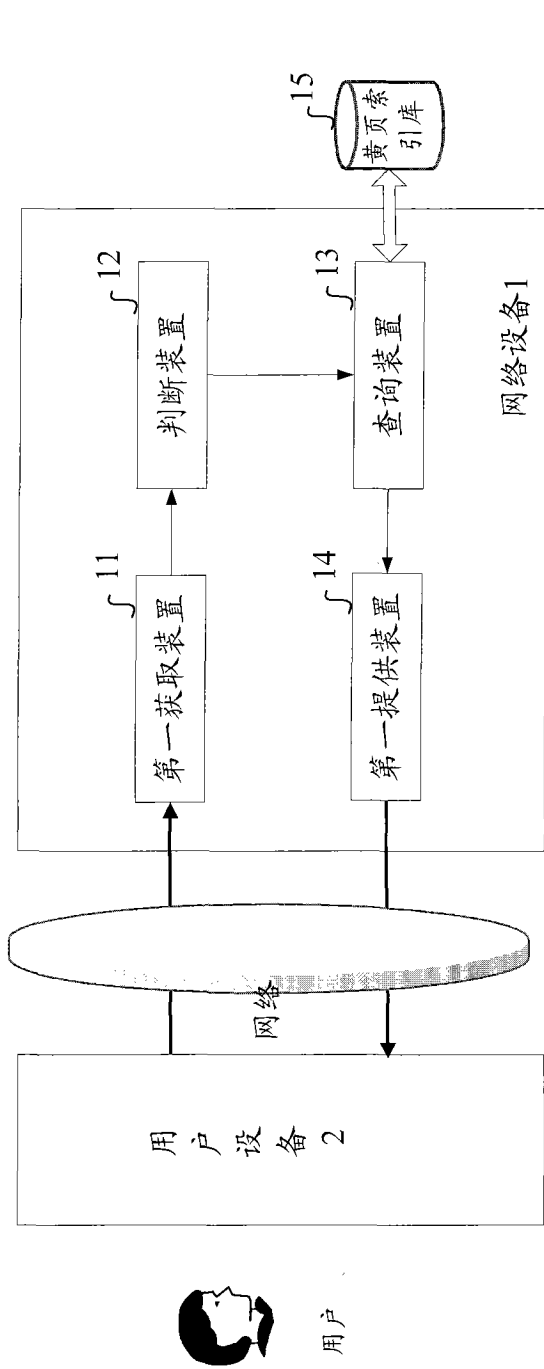


图 2

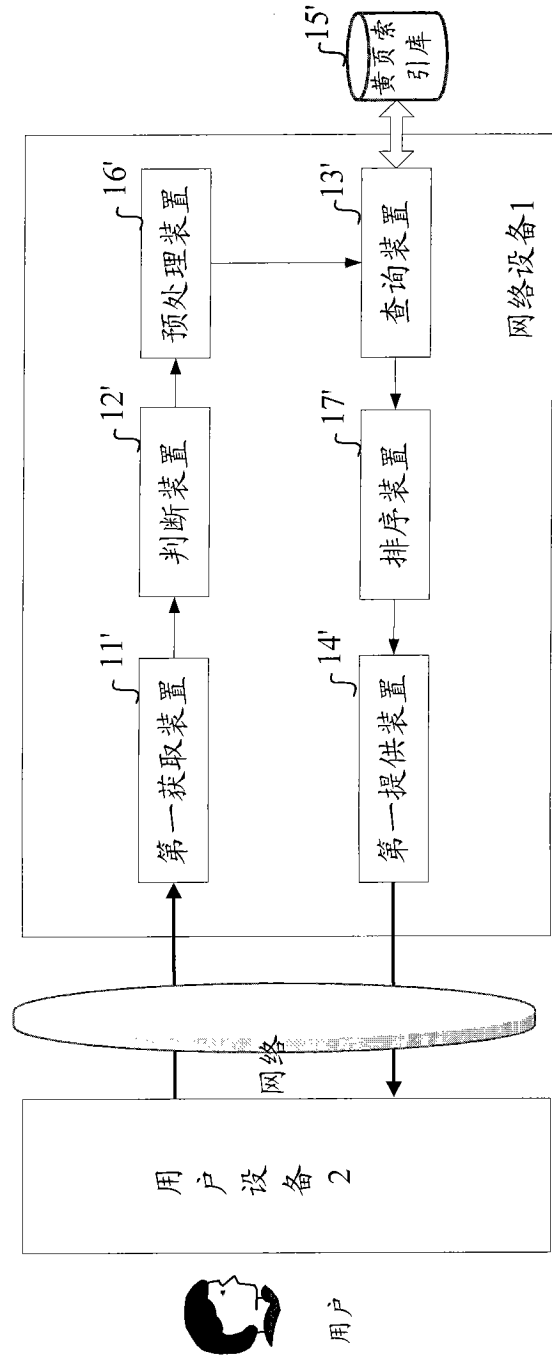


图 3

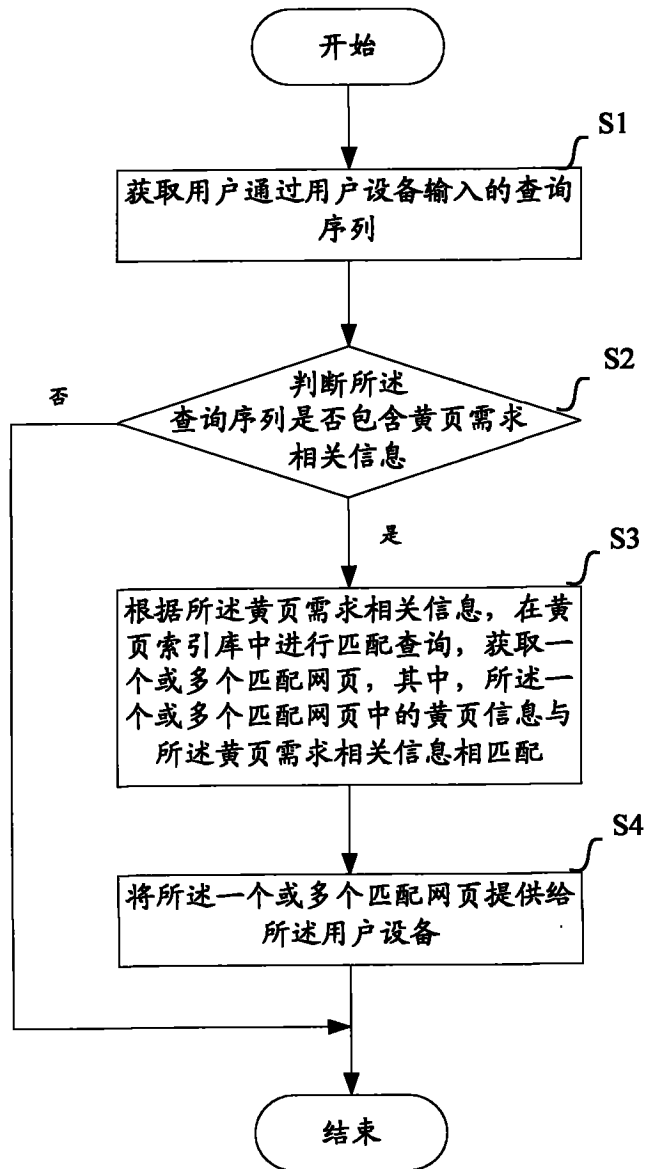


图 4

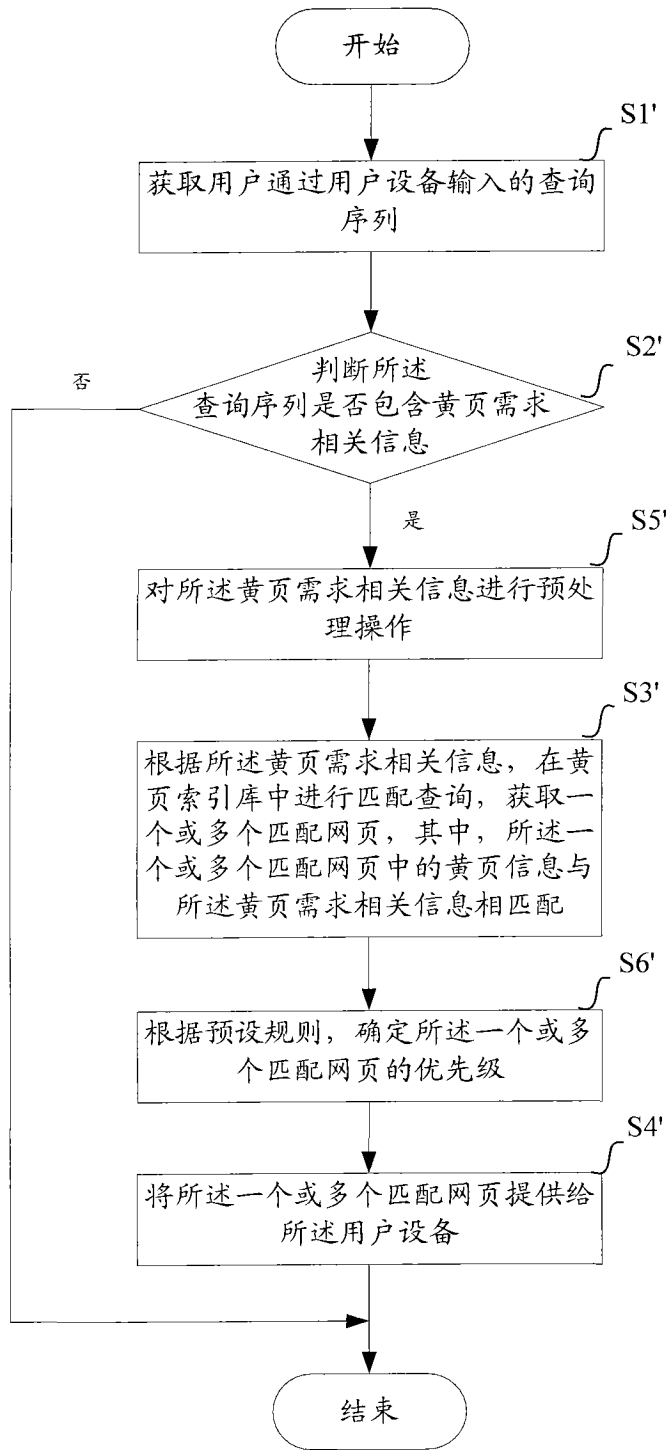


图 5